DOCUMENT RESUME

ED 217 297 CE 032 935

AUTHOR Angle, John
TITLE Learning and Earnings: The Loose Connection. Final Report.
INSTITUTION Arizona Univ., Tucson. Dept. of Sociology.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Sep 80
GRANT NIE-G-78-0006
NOTE 130p.

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS Achievement; Career Choice; *Education Work Relationship; *Employment Level; *Experiential Learning; *Income; *Occupational Mobility; On the Job Training; Postsecondary Education; Research Methodology; Research Problems; Secondary Education; *Work Experience

ABSTRACT

This research project set out to provide information to young people about which occupations would provide them with the most valuable experience, valuable being defined in terms of later earnings. The project attempted to measure the value of on-the-job learning in terms of future income and to identify sequences of occupations which would maximize the value of the young person's experience. Instead, the project found that the premise is not correct. Individual differences in on-the-job learning do not, on the average, result in differences in earnings. While work experience, measured as the length of time a person has worked, has a substantial positive impact on a person's earnings, the impact does not come from individual differences in job-acquired skills. Therefore, the project's first proposal for individualized simulations of young people's entry into the work force was not needed. The same advice can be given to all: Start work early in as highly paid an occupation as possible and work continuously at it; the connection between learning and earning is a loose one. (Research for the study was conducted through a literature review and longitudinal surveys. Ways were devised to use all available information in a longitudinal survey, and a method of measuring on-the-job learning was devised.) (Author/KC)

ED217297

LEARNING AND EARNINGS: THE LOOSE CONNECTION

FINAL REPORT: NATIONAL INSTITUTE OF EDUCATION RESEARCH GRANT #NIE-G-78-0006

by

John Angle

Department of Sociology

University of Arizona

September, 1980

CE032935

ABSTRACT

LEARNING AND EARNINGS: THE LOOSE CONNECTION

Final Report: National Institute of Education Research Grant

#NIE-G-78-0006

This research project set out to provide information to young people facing entry into full-time employment on which occupations open to them would provide them with the most valuable experience. Value was defined in terms of later earnings. It was assumed, following human capital theory, that what a person learns at work, like what a person learns in school, may increase marginal productivity. This project proposed to measure the value of on-the-job learning in terms of future income and identify sequences of occupations which would maximize the value of the young person's experience. Instead, this project found that the premise is not correct. Individual differences in on-the-job learning do not, on the average, result in differences in earnings. Work experience, measured as the length of time a person has worked, has a substantial positive impact on a person's earnings but it does not come from individual differences in job-acquired skills.

The starting point of the project was to investigate the relationship between in-school learning and earnings. This relationship would have to be controlled for in the examination of on-the-job learning and earnings. There is a large literature on this topic, but in most of the studies of the impact of education on earnings, education is measured rather crudely, by simply the number of years of schooling a person has completed. Use of this indicator assumes that people learn the same things at the same rate, an inaccurate assumption. Chapter 2 tests the effect of other indicators of learning in school on earnings and occupational prestige, and finds that the measurement of education by its duration explains most of the impact of education on occupational achievement. Only a person's major field in high school or college has any noticeable impact on occupational achievement independently of highest grade completed, and these effects are rather small. It appears that it is primarily the length of time one has spent in school, and secondarily one's major field, that impacts on one's earnings, not what one actually learned or did not learn, apart from the average of people staying in school for a given length of time and taking certain subjects. To make an analogy to canned goods whose quality varies from can to can: it's the label, not the contents, that affects the price. Chapter 3 investigates whether the choice of a person's major field in college can account for the gap between the earnings of college educated women and men. It cannot.

Chapter 4 introduces a technique for measuring work experience in a longitudinal survey, that is, one which re-interviews the same people over and over, but which does not find out their work histories from them. The technique involves interpolating what they do between the times they are interviewed from what they are doing at the time of the interview. Thus, if working at both consecutive interviews, they are credited with working the whole interval; if working at one interview but not the other, with working one-half the interval; if working at neither interview, with no time spent working. One can reconstruct a person's whole work history this way, but with error. Chapter 4 discusses how to handle this error.

Chapter 5 applies the technique of inferring work experience to finding out what impact it has on earnings. Chapter 5 finds that while work experience has a positive impact on earnings that if one subdivides a person's total work experience into the lengths of time a person has spent working at different levels of task complexity with people, data, and things, a measure of what a person is learning on-the-job, one finds that work experience at a high level of task complexity has about the same effect on wages as work experience at any other level, i.e., what one learns or does not learn has little or no impact on later wages. Rather, it must be some other mechanism -- perhaps seniority provisions in contracts or informal seniority preference, or the ability of more experienced people to pick the better paid jobs -- which accounts for why work experience is related to higher earnings. Learning may yet have something to do with earnings. But it is the average learning of people with a certain level of experience, not individual variations, which is recognized in a person's pay. Employers may recognize that a person with a given amount of experience is likely to have a given level of skill. It is as if the skill level were invisible but the amount of experience is visible. As with the connection between education and earnings, the connection between on-the-job learning and earnings is rather loose. Chapter 5 raises the question whether work experience accounts for the gap between the wages of young women and young men. Returns to the work experience of young women is somewhat discounted relative to that of young men but not enough to account for the gap between their wages. Rather it is the tendency of young men to receive higher wages as they age and young women not to that accounts for all the wage gap between the two.

ACKNOWLEDGEMENTS

I have to acknowledge a large debt of gratitude to Robert Stump, Senior Associate at the National Institute of Education, for his interest and help. His thoughtful questions and encouragement to examine my assumptions have substantially improved this research. Thanks also go to the Work and Education co-grantee discussion group in which I participated and dicussed this project. Patricia Taylor, Burke Grandjean, Marshall Pomer, Howard Mitchell, Diana Robertson, Ruth Prywes, John Grasso, Robert Dauffenbach, and Samuel Osherson all contributed to the project. Reporting periodically to this collegial group and receiving their opinions greatly facilitated research. I was assisted by Steven Steiber, now with the American Medical Association, and David Wissmann of the Social Sciences Department of Avila College, while they were graduate students in the Department of Sociology at the University of Arizona. Dr. Wissmann in particular invested considerable effort into the project.

TABLE OF CONTENTS

CHAPTER 1:

INTRODUCTION

This paper reports research done for the National Institute of Education (NIE) under grant #NIE-G-78-0006. The project's original and official title is "The Impact of Occupational Experience in Occupational Sequences: Analysis and Simulation." The project intended to do the necessary substantive and methodolgical work to make possible a proto-type computer simulation of young people's entry into the U.S. labor force which could be useful in advising young people on optimal strategies in choosing their initial jobs.

This project is a response to NIE's request for research into the area of learning for a career, not just for a single job. The research was called for because it was recognized by NIE that traditional vocational counseling has tended to advise students on the skills necessary for a single job, not the sequence of jobs which most people hold in their working lifetimes. The problem was how to advise young people for a sequence of jobs instead of a single one. An initial response to the question might be to try to find out what occupational sequences there are, but this task becomes increasingly difficult as one deals with finer and more specific job names, i.e., the more information one has the harder it is to proceed, an absurd predicament. Instead this project proposed describing jobs by their location on three dimensions: the maximum complexity with which a person in a job has to work with people, data, and things. These scales have been devised by the U.S. Department of Labor for the Dictionary of Occupational Titles. Once a job is designated by three scale scores and a wage rate instead of just a name, it is possible to apply a powerful tool of statistical inference, regression analysis, which is not applicable if one thinks of careers as sequences of job titles. This project was made possible by this reconceptualization of what it means to have a job.

This research proposed to find out whether what people learn by working early on in their careers acts like formal education to prepare them for better jobs later. If some kinds of work experience early in the career have a payoff later in the career, then this information ought to be used by guidance counselors to advise young people on their optimal strategy for entering the labor force. The reconceptualization of working as scores on the people, data, things scales and a wage rate permits viewing the relationship between learning and earnings and learning and occupational mobility as on-going. Heretofore, the typical models relating learning to earnings and occupational mobility conceptualized the relationship as occurring once, when the person left full-time schooling for full-time work. This project proposed to estimate the parameters of this on-going relationship and use them in a computer simulation of a young person's career, in which the implications of various job choices, among likely alternatives, could be examined, as well as the impact of discrimination and various scenarios for the abolition of discrimination. The project made the assumption, quite common among economists and sociologists, that the U.S. labor market responds to individual variations in a person's knowledge and skills. In fact, the whole point of the simulation was to advise young people to pick not just the highest pay rate available to them but the mix of pay and valuable opportunities for on-the-job learning which would maximize their lifetime earnings. This project has found that this assumption is not valid. Consequently, the simulation phase of the project has been abandoned. In a sense, it is a shame that the U.S. labor market does not work the way it was assumed to work. It is part of our national mythology that individual learning is recognized and rewarded. The fact that it is not and that individual knowledge and skill differences apart from the average of those with the same level of education and length of experience are not recognized, is a very important fact that future policy and research have to come to grips with.

This report is organized in the following way. The essential research questions faced at the beginning of the project are presented. Answering them and explaining how methodological problems were overcome by the invention of new techniques is the body of this report.

Essential Questions

This project proposed to analyze how what a young person learns by working results in occupational mobility and higher pay later in his or her career. To accomplish this objective one needs to be able to measure what people learn by working, not an immediately or obviously feasible goal at the beginning of the project, since the primary measure in use then was simply the length of time a person has worked under the assumption that people who have worked longer know more. This measure is not particularly informative. However, when one looks carefully at the other principal measure of learning used by sociologists and economists, the measurement of what a person has learned in formal education, by the number of years of schooling the person has completed, one sees that both these fields have been relatively unconcerned with the careful examination of what people learn from their experience either in school or on the job. The first essential research question that this project has to address is whether the measurement of learning in school by the length of the school experience, i.e., the number of years of formal schooling, is adequate. This project sought to find out ~~it~~ the labor market responds to individual variation in what people know because of their work experience. However, in order to isolate this effect one has to find out whether controlling for the number of years of schooling a person has completed is an adequate control for education, i.e., that one has isolated learning on the job from what the person has learned in school. Thus, before one can investigate what effect on-the-job learning has on later occupational mobility and wages, one has to find out whether years of school completed is an adequate measure of learning in formal education. Chapters 2 and 3 answer this question.

How does one go about measuring on-the-job learning? While the length of time a person has worked, usually measured in years, is the conventional indicator, it is possible, however, with the National Longitudinal Surveys of the Labor Market Experience of Young People, N.L.S. (cf. Center for Human Resource Research, 1976), to determine how long young women and men have worked in particular maximum levels of complexity with three aspects of task performance, work with people, data, and things. Thus, it is possible to divide up a young person's total length of time working into the lengths of time he or she has worked at different levels of complexity on each task dimension. This approach is an extension of the traditional indicator, total length of time worked, and its assumption that on-the-job learning is proportional to the length of time a person has spent working. Only in this instance, one can distinguish between work environments in which there is relatively little and relatively more to learn about working with people, data, and things. The N.L.S. surveys use the 1960 Census Occupational code which makes it possible to estimate the average people, data, things score of each occupation a person is identified as working in since there is a Current Population Survey of the same population in which occupations are double-coded, once in the 1960 Census code, once in the Dictionary of Occupational Titles code, for which there is a unique people, data, things scale score. This methodological innovation which makes this project feasible is described in chapters 4 and 5. The second essential research question is how on-the-job learning, measured as the lengths of time young people have worked at different levels of complexity with people, data, and things, affects later levels of complexity in working with people, data, and things, and wages. Chapter 5 deals with this question. Special attention is paid to the question of how the labor market works differently for young women than it does for young men. Gender has a very large impact on labor market experience, especially earnings, larger than color. Chapters 3 and 5 examine the issue of gender differential effects.

CHAPTER TWO:

EDUCATION INDICATORS AND OCCUPATIONAL ACHIEVEMENT

ABSTRACT OF CHAPTER TWO

Many sociological and economic studies assume that the variable, Highest Grade Completed in School, is by itself an adequate measure of people's education for the purpose of explaining their occupational achievement. Use of Highest Grade Completed as the sole measure of education has at least two major shortcomings. It assumes people have 1) learned the same amount of 2) the same thing in an academic year. Supplementary education indicators are identified and tested to see if they have a substantial impact on occupational prestige or earnings. Background social statuses and Highest Grade Completed are controlled for in this test. Only measures of subject matter studied in high school or college have a statistically significant relationship with occupational achievement net of Highest Grade Completed and social background variables. Highest Grade Completed is quite adequate in measuring the impact of education on occupational achievement without help from the measures of subject matter studied.

INTRODUCTION

In economics, Human Capital Theory provides an explanation for why people with more education earn more money than people with less education (Becker, 1964; Mincer, 1974). The explanation is an application of marginal productivity theory, under the assumptions that education and intellectual ability stand in a causal relationship to higher marginal productivity. In sociology, education is used as an explanation of the prestige level of a person's occupation as well as level of earnings. It has been observed that many sociological models of occupational prestige and earnings, particularly those in the "status attainment" tradition, have little theoretical underpinning (Coser, 1975; Burawoy, 1977). It is apparent though, considering the ease with which Human Capital Theory is incorporated in the status attainment literature (cf Stolzenberg, 1975), that most sociologists have assumed the truth of the marginal productivity explanation of the relationship between education and earnings. Such an assumption is not remarkable. Indeed, the notion that education raises productivity, thereby qualifying people for higher paying jobs, and incidentally, providing a legitimate basis for wage differentials, is virtually a cornerstone of American civic culture (Jencks et al., 1972).

This chapter re-examines the relationship between education indicators and occupational achievement, a single expression for the prestige of a person's occupation and earnings. Quite a bit of research has explained occupational achievement in terms of one indicator of education, Highest Grade Completed in School, or the similar measure, Number of Years of Schooling (cf *inter alia*, Becker and Chiswick, 1966; Blau and Duncan, 1967;

Haley, 1973; Featherman and Hauser, 1976). Such a measure of education is widely recognized as incomplete (Griffin and Alexander, 1978), but it is not obvious what other indicators of education are needed to supplement Highest Grade Completed in School as a measure of education.

This chapter uses simple models of the process of education and its impact on occupational achievement to identify supplementary indicators of education which are collected in large surveys of the labor force. The task of this chapter is to test whether any of these supplementary indicators of education explains some part of occupational achievement not explained by the traditional measure of education, Highest Grade Completed. This task is methodological. We are examining the adequacy of Highest Grade Completed as an indicator of education in the explanation of the impact of education on occupational achievement. In testing whether any of the supplementary indicators of education have some impact on occupational achievement net of Highest Grade Completed, we control for background social variables. We control for these variables since they might account for some of the zero-order relationship, if any, between a supplementary education indicator and occupational achievement. We are, of course, looking for the impact of the supplementary education indicator *per se* on occupational achievement so it is appropriate to make these controls. Figure 1 illustrates the relationships which are of interest to this investigation. They are the relationships between the supplementary indicators of education, represented by dotted lines, and occupational achievement.

(Figure 1 about here)

Highest Grade Completed in School

Most survey research on occupational achievement has adopted the usage of the U.S. Bureau of the Census in measuring education. The U.S. Bureau of the Census introduced a question on highest grade completed in the 1940 Census. "Grade" refers to an academic year (Shryock and Siegel, 1973: 328, 329). This measure is often referred to as "number of years of schooling," and as such can be extended to measure education at stages where the concept of "grade level" has no widely recognized meaning, i.e., graduate education beyond the master's degree or professional degree. Highest Grade Completed or Number of Years of Schooling are both measures of education by its cumulative duration. Highest Grade Completed has the virtue of being relatively objective and fairly easily recalled. As Duncan (1969:104) points out, such a measure may be correlated with other aspects of education, such as its quality. Since most educational institutions have some minimum standards for promotion, Highest Grade Completed is partially a measure of intellectual achievement. It is also correlated with intellectual ability (Griliches and Mason, 1972). As a variable, Highest Grade Completed is amenable to tabular analysis and regression procedures. It reduces a potentially complex multi-dimensional concept to a simple interval scale.

There is probably more to education's effect on a person than the length of time spent in an institution. Two problems with Highest Grade Completed as the sole measure of education are particularly acute. First, its use ignores the fact that different people learn different things in school. Since courses of study are more individualized and diverse (i.e., more electives, more tracks, more degree programs), the farther one is along in one's education, the poorer the assumption of the uniformity of what is learned. Secondly, the use of Highest Grade Completed as the sole measure of education assumes that everyone who completes a grade has learned the same amount of what is available to be learned. Ranges and variances in standardized achievement testing, as well as the distribution of letter grades, belie this assumption. There is, therefore, reason to measure the effect of other indicators of education on occupational achievement.

Models of Education and Its Impact on Occupational Achievement

Information on education besides the highest grade a person completed in school is collected in surveys. Often, however, such surveys are of a rather specialized population. The National Longitudinal Surveys of Labor Market Experience (N.L.S.) do not have this problem. They are surveys of large probability samples of birth cohorts of the U.S. population. This chapter takes the surveys of the two younger birth cohorts, men and women, aged 14 to 24 in 1966, and examines whether the supplementary education indicators can explain occupational achievement beyond what Highest Grade Completed can explain. Naturally, this test is made net of the effect of a block of background variables and individual attributes which might explain both the nature of a person's educational attainment and his or her occupational achievement.

We conceptualize the impact of education on occupational achievement to be a function of 1) what a person learns in school, and 2) the certification of that learning in the form of diplomas, degrees, certificates, and transcripts. It may be that the occupational achievement of young people once they leave school and go to work full-time is largely determined by the number of years of schooling they have completed and the degrees they have in hand, rather than by the knowledge they have acquired. Such a situation might be the case, if the gatekeepers to the labor force have no way of ferreting out an individual's knowledge, assuming that it has some relevance to marginal productivity. Young people coming out of school may be like canned goods, i.e., their contents are unknown and to be guessed at only from their labels -- the transcripts and diplomas which are awarded for staying in an institution for a given length of time. If this view of the relationship between education and occupational achievement is correct, then Highest Grade Completed might be as adequate a measure of education as will have relevance for explaining occupational achievement. However, the possibility is not precluded in this paper that the amount learned of a particular subject may make a difference in occupational achievement quite independently of a particular set of educational credentials.

What is learned in school is conceptualized as a function of 1) the subject taught, 2) the length of time a person is exposed to instruction, and 3) the quality of instruction. The rationale for this simple model of learning is that a student will learn a subject in proportion to the quality of the instruction and the length of time available for learning. The impact of education on occupational achievement is conceptualized as a joint function of what is learned and its certification. Equation #1 presents the relationship of education to occupational achievement in functional form. Equation #2 does the same for the relationship between what is learned and the survey indicators of education.

Occupational Achievement = fn (What is Learned) (Certification) eq. 1

What is Learned in School = fn (Subject) (Quality of Instruction) (Time) eq. 2

The literature on Subject Matter and Quality of Instruction is reviewed briefly below.

Subject Matter

There is a clear tendency for what is taught students to become more diverse at higher levels of schooling. The curriculum of the early elementary school grades is fairly uniformly focussed on basic literacy and numerical skills. However, by high school, students have sorted themselves out, or have been sorted out, into a number of quite distinct educational tracks, such as commercial, college preparatory; vocational or general studies, between which mobility may be at least somewhat restricted. In education beyond high school, there is yet more diversity and individual choice. Since tracks, major areas, or degree programs, in short, subject matter, tend to prepare students for particular vocational roles in many instances, choice of subject matter should be expected to affect occupational prestige and earnings later. Koch (1972) reports higher rates of return in earnings for majors in such areas as mathematics, accounting, economics, and psychology. Ashenfelter and Mooney (1968) show that field of graduate study explains more variance in earnings than the number of years that graduate study takes. Griffin and Alexander (1978), with data on almost 1000 male graduates from a sample of high schools found that high school track, measured by binary variables for college preparatory and vocational-commercial tracks, and the number of math or science courses taken, and college major, measured by binary variables for engineering and business, has a significant impact on occupational achievement. In particular, majoring in business or engineering in college added, on the average, more than $2,000 to the annual earnings of the men in their sample.

Quality of Instruction

Although quality of instruction is difficult to measure directly, it is possible to develop indicators of factors likely to be closely associated with quality of instruction. These are, for example, indicators of resources expended per student, the prestige of an educational institution (usually applicable only to colleges and universities), the perceived effectiveness of the institution, or the average level of achievement of students in the institution. All of these indicators describe the educational environment a typical student in the institution might encounter. Individual experiences may vary, of course.

It has been found that an indicator of school quality, such as expenditure per capita, is positively related to students' earnings at a later time (Welch, 1966). However, a great deal of research has shown that indicators of school quality are quite collinear in their relationship to later earnings and occupational prestige with indicators of background social status (Coleman, 1966; Astin, 1968; Bowles, 1972; Jencks et al., 1972). Thus, it is very difficult to sort out the effect of school quality on occupational achievement from those of background social statuses. Nevertheless, several recent studies, which have controlled for students' background social statuses, have found a net relationship between a measure of school quality and later occupational achievement. Using the Gourman (1967) Index of Institutional Quality, an average of subjective ratings of the components of post-secondary institutions, Wales (1973) found that those attending institutions in the upper fifth of the scale received higher returns to schooling than those who had not. Alwin (1974) reached much the same conclusion when he distinguished between "prestigious" and "non-prestigious" universities, using a variety of indicators of institutional quality.

The Data

The data for this examination of the relationship of education indicators to occupational achievement are the surveys of cohorts of young men and women conducted by the Center for Human Resource Research in the program of the National Longitudinal Surveys of Labor Market Experience (Center for Human Resource Research, 1976). There are 5,159 young women and 5,225 young men, aged 14 to 24 in 1966. They are a national probability sample of their birth cohort. The young women were interviewed in 1968, 1969, 1970, 1971, 1972, 1973, and 1975. The young men were interviewed in 1966, 1967, 1968, 1969, 1970, 1971, 1973, and 1975. The young men and women are thus 23 to 33 years of age at the time of the last available survey, that of 1975.

Whatever effect education will have on the occupational prestige and earnings of those in the labor force, it must be in a relatively early stage of these peoples' work lives. Such a restriction is an advantage since it tends to isolate the effect of education per se on occupational achievement from the effect of differential occupational experience. The well-educated may, on the average, tend to be placed in occupations where there is much to learn and thereby continue learning on the job, confounding the effect of occupational experience with that of education where learning opportunities on the job cannot be controlled for. Also, by looking at young people one avoids the issue of the obsolescence of education. This study is limited to people who have clearly begun at least to make the transition from full-time study to full-time work. Many Human Capital Theorists (cf Mincer, 1974) prefer a "commencement" model of work activity, in which the transition from education to work is assumed to be instantaneous. The "Commencement" model is a poor description of reality, but the need to draw an arbitrary line between those primarily engaged in education and those primarily engaged in work remains. This study uses the following criteria to make this distinction: at least 24 years of age, in the labor force under the Census definition, and working at least 30 hours a week in his or her current job. Observations, not people, are sampled. If a person does not meet the criteria for inclusion in one wave, he or she may in another wave. An observation which meets the criteria is selected from the observations on young women and men. There are 10,052 such observations on 3,437 men, and 3,954 such observations on 2,076 women. Observations with missing data are deleted. Observations on the same person over time are not independent. The "effective N" (Kish, 1965:162) of the sample for the purpose of hypothesis testing is conservatively taken to be the number of cases, not the number of observations, in an analysis. Cases are weighted by the reciprocal of the number of observations on them. Cases are also weighted by the inverse of the sample weights, that is, cases from over-sampled strata are constrained to be a proportionately smaller fraction of the cases entering the analysis and vice versa for cases from under-sampled strata.

The education indicators of the National Longitudinal Surveys of young people are given below. The sources of these variables are given in the codebooks of the National Longitudinal Surveys (Center for Human Resource Research, 1976). Where data are not drawn from an interview, the source is noted.

Supplementary Education Indicators

Subject

Type of high school curriculum of most recent high school year:

vocational
commercial
college preparatory
general
contrast category: people who never entered high school, people missing data on this variable

Field of study of last post-secondary degree received:

humanities
education
natural science
business
social science
contrast category: other fields, people who received college degree, people missing data on this variable

Time

Highest grade completed in school

Quality of Instruction (Primary and Secondary Education)

School district wide annual expenditure per pupil in average daily attendance, adjusted for local prices (cf Kohen, 1973)

A normalized index of school quality, which includes the following items:

per pupil availability of library facilites
pupils per full-time teachers
full-time equivalent counselors per 100 pupils
salary of a beginning teacher with a bachelor's degree and no experience, adjusted for local prices (cf Kohen, 1973)

Quality of Instruction (Post-secondary Education)

The ratio of students to faculty in most recently attended institution (calculated from full-time equivalents, from published information)

The ratio of expenditures to students in most recently attended institution (calculated from full-time equivalents, from published information)

Certification

- high school degree
- associate's degree
- bachelor's degree
- master's or doctoral degree
- contrast category: people with no degrees or other degrees, people missing data on this variable

Analysis

The question of interest is whether the supplementary indicators of education have an important effect on occupational achievement. In order to answer this question, a number of variables which are widely thought to affect both occupational achievement and the quantity and quality of education a person receives should be controlled for. Not controlling for such variables in an examination of the relationship between the supplementary education indicators and occupational achievement could lead to misleading overestimates of the importance of these indicators. Similarly, the effect of Highest Grade Completed on occupational achievement should be controlled for, so as not to mistakenly inflate the estimates of the relationships between the supplementary education indicators and occupational achievement.

What is the functional form of the relationship between occupational achievement, the education indicators, and background social variables? Simplicity of analysis recommends estimating a linear model. Equations #3 and #4 express occupational prestige and earnings in terms of a linear baseline model.

$$Y_1 = b_{10} + b_{11}X_1 + b_{12}X_2 + b_{13}X_3 + b_{14}X_4 + b_{15}X_5 + e_1 \qquad \text{eq. 3}$$

$$Y_2 = b_{20} + b_{2y}Y_1 + b_{21}X_1 + b_{22}X_2 + b_{23}X_3 + b_{24}X_4 + b_{25}X_5 + e_2 \qquad \text{eq. 4}$$

where,

Y_1 = Occupational Prestige (Duncan socioeconomic index score)

Y_2 = Annual Earnings in Previous Year (inflated to 1975 dollars)

X_1 = Age

X_2 = Gender (1=female, 0=male)

X_3 = Color (1=black, 0=other)

X_4 = Occupational Prestige (Duncan score) of Father or Head when Respondent was 14

X_5 = Highest Grade Completed in School.

IQ score was considered as a possible control variable. It is measured in the National Longitudinal Surveys. However, it has a very high rate of missing data, and, worse, the likelihood of a case missing data is closely correlated with at least two of the control variables, Gender and Age. All regression equations estimated in the course of this research were also estimated with IQ score (actually the decile of a person's score on any of several tests of intellectual ability) as a control variable. Previous research has suggested that failure to control for IQ may lead to inflated estimates of returns to education (cf Griliches and Mason, 1972; Griffin, 1976). Such is the case in the regression of earnings on Highest Grade Completed in School but is not necessarily the case in the regression of earnings on other education indicators or the regressions with occupational prestige as a dependent variable. However tantalizing these variations may seem, there are too many missing data on IQ scores in this data set to reliably arrive at any conclusion about the returns to an education indicator net of IQ. Inclusion of IQ score does not alter the sign or statistical significance of a supplementary education indicator, net of the other control variables, in any regression reported in this paper.

Figure #1 illustrates the tests to be performed. Table #1 gives the estimated coefficients of equations #3 and #4. The Durbin-Watson test indicates significant autocorrelated error in every equation. Accordingly, an Orcutt transformation is performed. Equation #5 illustrates the procedure. If the e_t's of $Y_t = b_0 + b_1X_t + e_t$ are found to be correlated over time, then an Orcutt transformation puts this equation in the form:

$$Y_t - pY_{t-1} = b_0^* + b_1(X_t - \hat{p}X_{t-1}) + v_t \qquad \text{eq. 5}$$

where,

p = the estimated coefficient of autocorrelation of the e_t's

$b_0^* = b_0(1-\hat{p})$

$v_t = e_t - \hat{p}\, e_{t-1}$

and it is expected that $Cov(v_t, v_{t-1}) = 0$. The parameters to be estimated are the same. However, it is expected that the transformed equation from which they are estimated will have negligible autocorrelated error.

(Table 1 about here)

The strategy of the analysis is to add each one of the supplementary indicators of education to equations #3 and #4 and to compare the resulting increment in explained variance with the increment in explained variance when the variable, Highest Grade Completed, is added to the equations. If there is a substantial increment in the explained variance of occupational achievement with the addition of the supplementary education indicators, then these ought to be brought into models of occupational achievement.

Findings

All the supplementary education indicators available in the N.L.S. of young people were added, one at a time, to the baseline model of occupational achievement and their significance tested. Only two of the indicators have statistically significant relationships to occupational achievement net of the baseline model. Both of these indicators, Type of High School Curriculum and Field of Study of Last Post-secondary Degree (College Major), are indicators of what people studied in school. Neither variable, however, results in a major increase in the explained variance of Occupational Prestige or Earnings.

Table #2 shows that subject matter studied in school has a somewhat more important relationship with earnings than with occupational prestige. Taking a college preparatory track in high school yields a rather large return in earnings and some return in occupational prestige but not greater than that of the commercial track in terms of prestige. The commercial track in high school is the second most remunerative and has the largest positive effect on occupational prestige. Majoring in business in college yields the largest increase in earnings for any college major, $1,777 a year and also the biggest increase in occupational prestige. Majoring in natural science yields the next largest increase in earnings, $603 a year, but no increase in occupational prestige. There are a number of points of agreement between these findings and those of Griffin and Alexander (1978) but a number of the high school track and college major effects they found not to be significant are significant here.

(Table 2 about here)

Conclusions

This chapter's intent is methodological. It raises the question of whether Highest Grade Completed is an adequate measure of education for use in models of occupational achievement. Highest Grade Completed in School is an enormous simplification of what one might suppose are the various dimensions of education which could affect occupational achievement. At the very least, exclusive use of Highest Grade Completed as a measure of education makes two implausible assumptions: 1) that people learn the same thing at 2) the same rate. This chapter has conceptualized education as the amount a person has learned about different subjects. The data set on which this paper is based, the National Longitudinal Surveys, does not measure directly what the young people surveyed know. However, there is some approximate information on what subjects they have taken in school and the quality of the educations they received, what degrees they have received, as well as information on the number of years of schooling they have completed. These education indicators can be presumed to indicate how much people have learned of what subjects in school.

This chapter has tested whether any of the variables hypothesized to be indicators of what people learn in school have an effect on occupational achievement, net of background social variables (known to be factors in occupational achievement independent of education), and net of the widely used measure of education, Highest Grade Completed. If any of the supplementary education indicators make an impact on occupational achievement independently of Highest Grade Completed, this finding would be evidence that there is something in the content of what is learned in school which has an impact on occupational achievement beyond that of the number of years of school one has completed or the educational credentials one holds. It turns out that only indicators of the subject matter a person has studied in school have an effect on occupational achievement beyond that of the number of years of schooling completed, and the effect of subject matter taken in school on occupational achievement, net of Highest Grade Completed, is small. This finding settles the methodological question the paper raised. Highest Grade Completed stands very well by itself as the sole indicator of education in models of occupational achievement. It may be supplemented by indicators of what subject matters people have studied in school, but this refinement is optional because these indicators do not have a large effect on occupational achievement net of Highest Grade Completed.

The substantive sociological questions this finding raises, however, are not settled here. Indeed, they are complex and the data set used in this chapter is not at all useful in settling them. Why does Highest Grade Completed explain so much of the impact of education on occupational achievement? It may be that occupational achievement really is a function of what a person has learned in school, rather than a person's number of years of schooling, but that the number of years of schooling indicates how much a person knows better than the various education indicators used in this paper. Duncan (1969:104) has essentially taken this position. The alternative possibility is that the quantity of information a person has learned in school counts for very little in terms of occupational achievement in and of itself. Rather, in this alternative, it is the final level of schooling one has attained and the educational credentials one has received to certify this attainment which make the important difference with employers, bosses, personnel committees, and clients, those who keep the gates to occupational achievement. The failure of most of the supplementary education indicators to affect occupational achievement at all suggests that this latter alternative -- that it is the credential rather than variations in individual knowledge which better explains differences in occupational achievement -- is correct.

However, the theoretical questions are still moot. A careful longitudinal study of a birth cohort as it passes through a school system and on to higher and/or vocational education and on into the labor force is required to find out whether how much one knows is an important determinant of one's final level attained in formal education and whether knowledge or educational credentials make the large difference in occupational achievement. This study would need to measure people's knowledge at each stage of their education and labor force participation. Although the National Longitudinal Surveys are well provided with data on schooling as large surveys of labor force participation go, they do not permit the testing of theories of why people complete a given number of years of schooling, or whether the amount of information a person has learned has much to do with his or her final level in formal education. Too many important variables are missing, such as standardized achievement test scores, or, as with the case of IQ scores, too many data are missing, or the time order of events in a person's education cannot be reconstructed. This chapter has established that a person's number of years of formal schooling, his or her final level attained in formal education, is quite adequate by itself to measure the impact of education on occupational achievement. This chapter must remain agnostic on the theoretical explanations of this finding.

Table 1. Least Squares Regressions. Estimates of Coefficients of Equations #3 and #4 for Young Women and Men, Aged 24-33, in Labor Force and Working More than 30 Hours a Week[a]

	Equation #3 Occupational Prestige	Equation #4 Annual Earnings[c]
	Unstandardized Coefficients	
Occupational Prestige[b]	*	$29.99
Age[d]	-.16@	211.59
Gender[e]	3.42	-5,755.32
Color[f]	-6.67	-1,770.94
Occupational Prestige of Father or Head of Household when Respondent was 14[g]	.09	17.08
Highest Grade Completed in School[h]	4.72	613.03
Constant[i]	-11.50	-8,176.61
Coefficient of Autocorrelation	.64	.62
R^2	.135	.115
N[a]	4,188	4,178

[a]All observations are weighted by the reciprocal of the number of observations on an individual appearing in the sample, making the effective N the number of cases rather than the number of observations in the analysis. Observations are also weighted by the inverse of the probability of an individual's becoming a respondent in the survey.

Coefficients in this table have been estimated from Orcutt transformed variables. See equation #5 in the text.

Coefficients are statistically significant at the .05 level according to an F-test, unless marked by a @.

[b]Occupational Prestige is measured by the Duncan Socioeconomic Index (Duncan, 1961).

Table 1 -- continued

[c] The sum of wages and salary in previous year, and earnings from a farm or business in previous year. This latter variable could take on negative values. If data were missing on one variable, its value was assumed to be zero. If data were missing on both variables, a missing data code was assigned to their sum. All dollar figures from years before 1975 have been inflated to 1975 dollars.

[d] Age in years.

[e] Gender is a binary variable, which takes on the value 1.0 if a person is female, 0.0 if male.

[f] Color is a binary variable, which takes on the value of 1.0 if a person is classified black, 0.0 otherwise.

[g] Occupational Prestige of Head is measured by the Duncan Socioeconomic Index (Duncan, 1961).

[h] Highest Grade Completed in School measures whether a person completed a school year from the first grade through the end of four-year college. Beyond that point it is a measure of the number of school years spent in an educational institution.

[i] This quantity equals $b_0^*/(1 - \hat{\rho})$ where b_0^* is the intercept of the regression of Orcutt transformed variables and $\hat{\rho}$ is the coefficient of autocorrelation.

Source: National Longitudinal Surveys of Labor Market Experience, Cohorts of Young Men and Women (Center for Human Resource Research, 1976).

Table 2. Unstandardized Coefficients of Supplementary Indicators When Added to the Baseline Models of Occupational Achievement, Equations #3 and #4, Young Women and Men, Aged 24-33, in Labor Force and Working More than 30 Hours a Week[a]

	Equation #3 Occupational Prestige	Equation #4 Annual Earnings
Type of High School Curriculum of most Recent High School Year		
vocational	-6.69	$529.78
commerical	5.66	842.07
college prepartory	4.78	2,134.32
general	-3.60	526.98
never in high school or missing data	0.0	0.00
increment in r-square	.00886	.00292
ratio of increment in r-square with this variable to increment in r-square when Highest Grade Completed in School was added to equation	.09417	.17186
Field of Study of Last Post-Secondary Degree Received		
humanities	2.19	$253.88
education	6.66	-244.39
natural science	-.90	603.38
business	13.04	1,777.40
social science	3.65	203.54
other fields, people who did not attend a post-secondary institution, and people missing data	0.0	0.00
increment in r-square	.00873	.0019
ratio of increment in r-square with this variable to increment in r-square when Highest Grade Completed in School was added to equation	.09888	.11183

Table 2 -- continued

[a]Education indicators whose net relationship to a dependent variable were not statistically significant at the .05 level with an F-test are omitted from table. F-tests for sets of binary variables, such as those for high school curriculum or college major, follow Kmenta (1971:371)

$$F = \left[\frac{SSR_Q - SSR_K}{SSE_Q}\right]\left[\frac{n - Q}{Q - K}\right],$$

where,

K = the number of explanatory variables before the test

Q = the number of explanatory variables after the new set of explanatory variables has been added to the regression.

All estimates from Orcutt transformed equations.

Source: National Longitudinal Surveys of Labor Market Experience, Cohorts of Young Men and Women (Center for Human Resource Research, 1976).

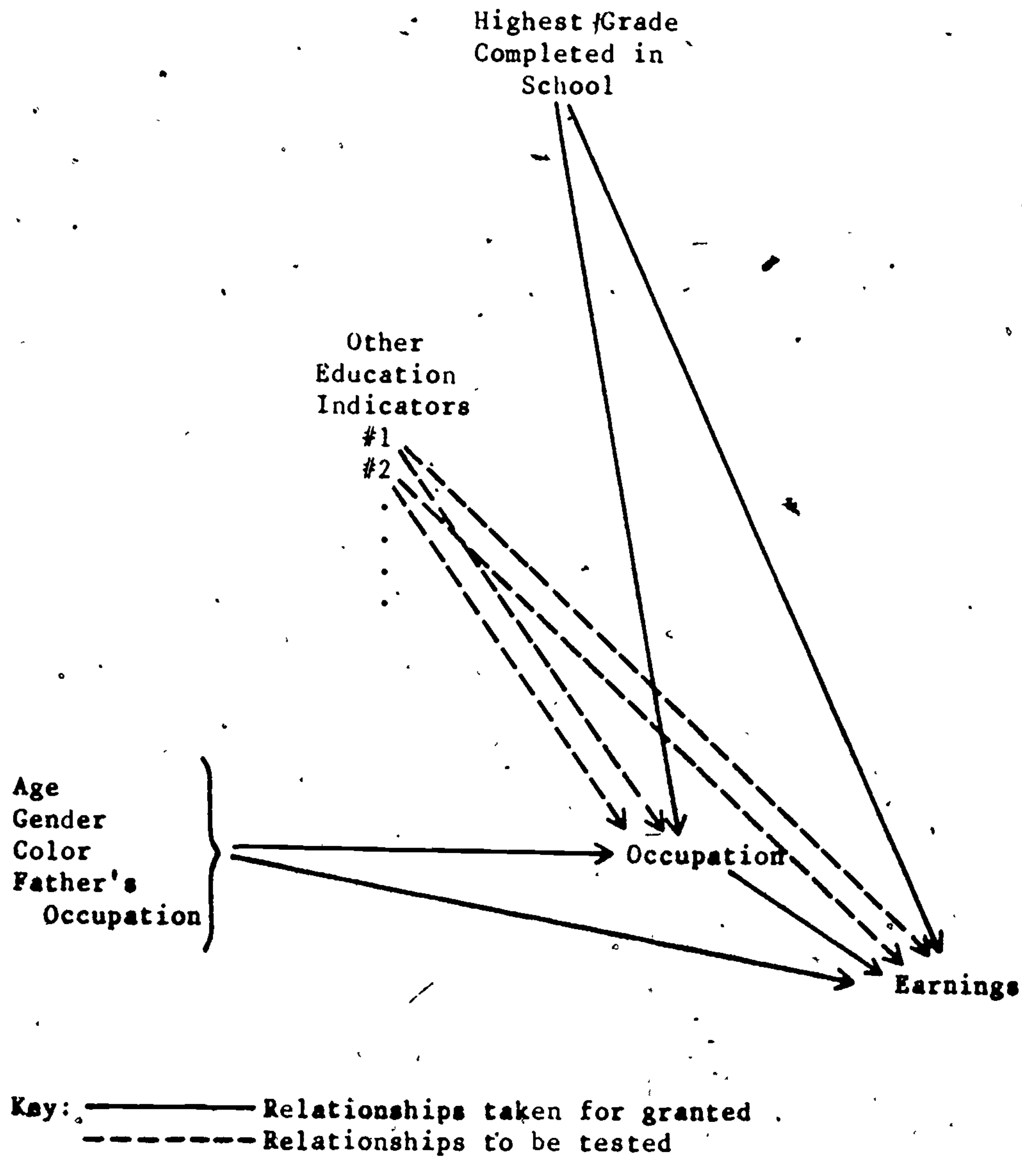

Figure 1. Schematic Representation of Test.

CHAPTER THREE:

GENDER, COLLEGE MAJOR, AND EARNINGS

ABSTRACT OF CHAPTER THREE

Studies that examine the gap between women's and men's earnings and the gap in their returns to education have used a person's years of school completed as the measure of his or her education. It may be that these gaps are produced by what subject matters men and women study rather than by discrimination. This chapter tests the effect of a person's major field in post-secondary education on his or her hourly wage to see if the content of what is learned in college, as opposed to the duration of the educational experience, can explain any of the gap between men's and women's earnings due directly to gender or any of the gap in returns to education. Data are taken from the National Longitudinal Surveys of the Labor Market Experience of young women and men. It is found that the direct effect of gender on the earnings of people with at least some college education is large, even net of a number of important control variables, and that controlling for major field of study reduces this gap only slightly. It is unexpectedly found that young women's return to a year of post-secondary education is higher than young men's, although not much so and not enough to offset the massive negative effect of being female on earnings.

Chapter 2 has shown that a person's number of years of schooling completed measures just about all there is to measure of the impact of education on earnings and occupational prestige. However, the subject one studies in school does have a measureable, if small, effect on occupational achievement. This chapter explores the possibility that, among people with some college education, the area of one's studies, one's major field, explains the gap between the wages of women and men.

Many studies have found that women earn less than men and that the rate of return to a year of education is lower for women than men. Both findings have been interpreted as evidence of discrimination against women in the labor force. However, it is possible that both the gap between men's and women's earnings and the gap between their rates of return to a year of education could be closed if the content of their respective educations could be measured. It might be that what women learn yields, in a non-discriminatory labor market, a lower rate of return than what men learn. To test this hypothesis we examine returns to the educations of people who have had at least some post-secondary education. Subject matters of study at this level of education are much more distinctive than at lower levels, since an individual's choice is much more free in course and major area selection. Differences between the genders in these choices may account for the earnings and the educational returns gaps. This chapter tests whether they do.

Specifically, this chapter reviews the research on differential returns to the educations of women and men, explanations for the gap, and problems with the use of the number of years spent in school as the sole indicator of education. A is test performed to see whether the major fields people choose in college affect the gap between the earnings of men and women and the rate of return in earnings to a year of education for men and women. The question of whether men and women majoring in the same areas have the same level of return in earnings to that choice is also examined. This test is conducted on young adults in the labor force, people who have not had much opportunity to develop experience working, a variable some have theorized to account for much of the gap between women's and men's earnings (Cohen, 1971; Malkiel and Malkiel, 1973; Featherman and Hauser, 1976). The dependent variable in this study is the natural logarithm of a person's hourly rate of pay. This variable is preferable to annual earnings since it has been pointed out that women may not work as many hours as men in any year (Oppenheimer, 1970; Fuchs, 1971). A logarithmic transformation of the dependent variable is both theoretically (Mincer, 1974:11) and statistically desirable (Stolzenberg, 1975:651,652).

Differential Returns to the Educations of Women and Men

The majority of studies on the subject have found that men convert their educations into earnings at a more favorable rate than women (Renshaw, 1960; Hines et al., 1970; Malkiel and Malkiel, 1973; Suter and Miller, 1973; Featherman and Hauser, 1976; King, 1977). Although there is ample evidence that females learn more than males in elementary school (Weitzman, 1975), the subject matters studied by the two genders become more distinctive at higher levels of education (Roby, 1975). The point is often made that women, for example, tend to avoid mathematics and since the study of mathematics is a prerequisite for entry into many well-paid technical and scientific careers, women may earn less because of their choice of subjects in school (Kagan, 1964). It has also been suggested (Becker, 1975:179) that women go to college partly "to increase the probability of marrying a more desirable man," and do not pay as much attention to learning, on the average, as men do in college.

Two studies, however, do not concur that men always have higher returns to a grade completed in school than do women. Cohen (1971) found in a survey conducted in Michigan that the returns in earnings to a year of education were the same for men and women. Mincer and Polacheck (1974) compared the hourly pay of single women and single men and found that single women had a higher rate of return to a year of education. However, married men had a higher rate of return in hourly pay to a year of education than married women. It is safe to say that the great weight of the literature points to men having a higher rate of return in earnings to a grade completed in school than women.

Years of Schooling

All of the research on differential returns to the educations of men and women which we were able to find in the literature have used the variable Highest Grade Completed in School as the sole measure of education. The U. S. Bureau of the Census introduced a question on "highest grade completed" in the 1940 Census. "Grade" refers to an academic year in school (Shryock and Siegel, 1973:328, 329). This measure is often referred to as "number of years of schooling," and as such can be extended to measure education at stages where the concept of grade level has no meaning, i.e., graduate education beyond the level of a master's or professional degree. It is used in this way in this paper, as the number of years of enrollment in post-secondary education. Highest Grade Completed, years of schooling, or years of enrollment are all measures of education by its cumulative duration. These measures have the virtue of being objective and fairly easily recalled. As Duncan (1969:104) points out, such a measure may be correlated with other aspects of education such as its quality. Since most educational institutions have some minimum standards for promotion, years of schooling are partially a measure of intellectual achievement. This measure is also correlated with intellectual ability (Griliches and Mason, 1972). As a variable, years of schooling is amenable to tabular analysis and regression procedures. It reduces a complex concept, education, to a simple interval scale.

However, it is reasonable to speculate that there should be more to education's effect on later earnings than just that of the length of time the person spends in an educational institution. Two problems with years of schooling as a sole measure of education stand out. First, its use as a sole indicator ignores the fact that people learn different things. Since courses of study are more individualized and diverse (i.e., more electives, more tracks, more degree programs) the farther one is along in one's education, the poorer the assumption of the uniformity of what is learned for the more educated. Secondly, the use of years of schooling as the sole measure of education assumes that everyone who completed a grade or a year of schooling has learned the same amount of what was available to be learned.

Neglect of the first problem, the assumption of the learning of the same thing in a year, is particularly acute in post-secondary education. By then people have sorted themselves, or have been sorted, into a variety of degree programs known to have important consequences on later earnings. Ashenfelter and Mooney (1968) have shown that field of graduate (post-baccalaureate) study explains more of one's later earnings than the number of years one puts into that area of study. Koch (1972) reports higher rates of return in earnings for college majors in mathematicized disciplines. Griffin and Alexander (1978), with data on almost 1,000 male graduates from a sample of high schools found that the number of mathematics or natural science courses a person takes, and one's college major have a significant impact on occupational achievement. In particular, majoring in business or engineering in college added, on the average, more than $2,000 to the annual earnings of the men in their sample. There clearly then is reason to ask whether taking field of study into account could change findings about discrimination against college educated women in the labor force.

The Data

The data for this examination of differential returns in earnings to education are the surveys of cohorts of young women and men conducted by the Center for Human Resource Research (1976) in the program of the National Longitudinal Surveys of Labor Market Experience. The young men and women were 14 to 24 years old in 1966. There are 5,225 young men and 5,159 young women. They are a national probability sample of their birth cohort. The young women were interviewed in 1968, 1969, 1970, 1971, 1972, 1973, and 1975. The young men were interviewed in 1966, 1967, 1968, 1969, 1970, 1971, 1973, and 1975. The young men and women are thus 23 to 33 years of age at the time of the last available survey, that of 1975.

By looking at younger workers one avoids the issue of the obsolescence of education. The study is limited to people who have clearly begun at least to make the transition from full-time study to full-time work. Many economists prefer a "Commencement" model of work activity in which the transition from education to work is assumed to be instantaneous (Mincer, 1974). The "Commencement" model is a poor description of reality, but the need to draw an arbitrary line between those primarily engaged in education and those primarily engaged in work remains. This study uses the following criteria to make this distinction: the person has enrolled in an educational institution after high school graduation, is at least 24 years of age, in the labor force under the Census definition, working at least 30 hours a week in his or her current job, and not missing data on a variable which enters the analysis. Observations on people making 35¢ or less an hour are discarded since the logarithms of such values would distort estimated coefficients.

Observations, not people are sampled. If a person does not meet the criteria for inclusion in one wave, he or she may in another. Table 1 displays the numbers of observations on women and men which enter the analysis from each year of the study. There are 1,677 observations on 764 young women and 2,831 observations on 1,359 young men. There are no observations on men in 1973 since information on hourly pay was not collected, and similarly, there are no observations on young men in 1966 because information on hours worked per week was not collected then. Since the women and men of this study are relatively young workers, 14 to 24 in 1966, it is not surprising that most of the observations on their labor force characteristics come from the latter waves of the study, when more of them have finished schooling and taken on adult responsibilities. It is possible that there are autocorrelated disturbances in a regression equation estimated with this data set. However, computing a coefficient of autocorrelation requires the loss of a massive proportion of the data set. Computing a coefficient of autocorrelation requires an estimate of a disturbance in the regression equation in waves t and t-1. Thus all observations on young men in 1975 would be lost since the t-1st wave is missing data on hourly pay. Also, computation of the coefficient of autocorrelation at t and t-1 requires that both observations meet the criteria for selection into the analysis and that neither is missing an observation. Whatever gain in efficiency of the estimates from using the coefficient of autocorrelation in an Orcutt transformation on the data set would not be justified by the biasing of the data set through the loss of observations.

Observations on the same person over time are not independent. The "effective N" (Kish, 1965:162) of the sample for the purpose of hypothesis testing is conservatively taken to be the number of cases, not the number of observations, in an analysis. Cases are weighted by the reciprocal of the number of observations on them. Cases are also weighted by the inverse of the sample weights, that is, cases from over-sampled strata are constrained to be a proportionately smaller fraction of the cases entering the analysis and vice versa for cases from under-sampled strata. See the footnotes of Table 3 for the formula used. The weighting technique used here permits the maximum use of all available information while permitting the use of statistics which assume simple random sampling in which the number of independent contributions of information is the same as the number of cases observed.

(Table 1 about here)

Date on the college major of young people in the National Longitudinal Surveys are given below in Table 2. Data on majors identified this way were collected in every wave. The question on college major in the first wave inquired about the "field of study" of the person's most recent college degree, if any. Later waves inquired about the field of study of the person's college degree since the last interview, if any. The unit of analysis of Table 2 is the observation, not the case. A person may change field of study between degrees, so weighting the observations to add to the number of cases would be misleading. As a glance at Table 2 shows, women tend to choose degree programs in primary and secondary education and in the humanities. Most of all, by comparison to men, women tend not to finish degree programs, at least by the time they are 24 to 33 years old. Men tend to take degrees in quantitative, technical, or scientific fields and in business. Women and men are about equally interested in the social sciences.

Analysis

We want to find out: 1) whether the choice of a field of study in college affects the gap between women's and men's earnings, 2) whether it affects the gap in returns to a year of education between women and men, and 3) whether women and men have different returns to the same field of study. The strategy used here to measure these effects is first of all to construct a baseline model of log hourly wage and then to test for the significance of the effect on log hourly wage of field of study and its interaction with gender. As these variables are added to the model of log hourly wage, the coefficients of Years of Enrollment in Post-Secondary Education and its interaction with gender are examined for change.

Estimated coefficients will depend on what variables are included in the baseline model. At first thought, it might seem appropriate to have a very inclusive baseline model, that is, to make as many controls as possible. There are three constraints however on the inclusion of variables into the baseline model. First, there is the practical constraint that if one follows the conventional practice of discarding a case that is missing an observation on a variable in the analysis, as we do, then the more variables one includes in one's regression, the fewer cases there will be and the more the sample will be biased by whatever factors determine which cases will have missing data. Secondly, there is the constraint that we do not want to make our baseline idiosyncratic. We have, for example, followed the practice of Featherman and Hauser (1978) and Jencks (1979) in measuring post-secondary education by its length in years separately from the number of years in primary and secondary education, and by including binary variables for post-secondary degrees received. The other variables in the baseline model are commonly controlled for in earnings determination equations. Thirdly, we do not want to control for variables which themselves may be causally posterior to a person's choice of field of study in college.

We control for age, color, the prestige score of the father's (or head of household's) occupations when the respondent was 14 (a measure of background social class); years of enrollment in post-secondary education, and for post-secondary degrees received. Gender is in the baseline model as are its interactions with years of post-secondary enrollment and post-secondary degrees. We want to see if controlling for major field affects the coefficients of gender and its interactions with years of post-secondary enrollment and degrees, coefficients which have been interpreted as indicative of discrimination against women and their education attainments. Equation 1 is the baseline model:

$$Y_{1t} = b_0 + b_1X_{1t} + b_2X_{2t} + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_{6t} + b_7X_{7t} + b_8X_{8t}$$
$$+ b_9(X_{2t}\cdot X_5) + b_{10}(X_{6t}\cdot X_5)\ f\ b_{11}(X_{7t}\cdot X_5) + b_{12}(X_{8t}\cdot X_5) + e_1 \quad \text{eq.1}$$

where,

Y_t = the natural logarithm of a person's hourly wage

X_{1t} = age in years

X_{2t} = years of enrollment in a post-secondary education

X_3 = color (1=black, 0=other)

X_4 = prestige score of father's (or head of household's) occupation when respondent was 14

X_5 = gender (1=female, 0=male).

X_{6t} = associate of arts degree (1=has one, 0=does not)

X_{7t} = bachelor's degree (1=has one, 0=does not)

X_{8t} = master's degree or doctorage (1=has one, 0=does not)

We would have liked to include a measure of intellectual ability, as a control variable. The N.L.S. has such a variable; a person's decile score on a variety of tests of intellectual ability. Unfortunately, more than half the cases are missing data on this variable. There are patterns in the missing data as well. For example, many more men than women are missing data on this variable. When the N.L.S. intellectual ability variable is used, it is with apology for these problems (Griliches, 1976). We have decided not to include it.

Findings

The coefficients of Table 3 are estimated from a semi-logarithmic equation, $\ln Y = b_0 + b_1X_1$ ....., which can be reexpressed as $Y = e^{b_0 + b_1X_1 + \ldots}$ ...... The former expression is perferable for purposes of coefficient estimation, while the latter is preferable for the interpretation of coefficients, since we would rather talk in terms of dollars than log dollars. Let's look at the effect on hourly earnings of being female in the baseline model. Net of the control variables, young women earn less than young men. Being female has a coefficient of -.4083. This number means that since $e^{-.4083} = .6648$, young women with at least some college education earn two-thirds of what young men with at least some college education earn, controlling for the variables of age, years of post-secondary education, degrees held, color, prestige of father's occupation, and the interaction between gender and years of post-secondary education, and the interactions between gender and degrees held. A thirty-three percentage point difference in hourly wages between college educated people who are quite similar in many respects relevant to their earnings potential, except for their sex, is enormous. It is doubtful that such a large difference can be explained away by differences in what women and men study in school. When one considers that we have ignored altogether the impact of gender on whether a person finds work at all, it is clear that gender is by far the largest factor affecting a person's ability to earn money by working.

It was expected that women would be severely penalized because of their gender. It was not expected that women would have a higher rate of return to a year of post-secondary education than men. In these data, women have a .5 percent higher rate of return to a year of post-secondary education than men. The coefficient is .0363, equivalent to a rate of return to a year of post-secondary education 3.7 percent higher than that of men. It should be pointed out that this advantage of women is much smaller than the direct disadvantage of being female. The finding that women have a higher rate of return to a year of education is quite contrary to what existing theories of ethnic discrimination in the labor force would lead one to suspect about gender discrimination. One would expect discrimination to lead to a discounting of the discriminated group's educational credentials. However, Oppenheimer (1970:99,100,114) has pointed out several facets of the female labor market which are consistent with the finding that women might have a higher rate of return to a year of education than men. These are that occupations typed female are characterized by a labor force which acquires its skills in school rather than on the job. Also, Oppenheimer sees some occupations typed female as having high educational requirements but because of occupational segregation by gender low pay. Either aspect of the structure of a female labor market might explain the finding.

Could it be though that the finding is an artifact? Could the Gender-Years of Post-secondary Education interaction be so closely correlated with Gender that we should look askance at its coefficient, even though it is statitically significant? The correlation between this interaction and Gender is .80, close but not unexpectedly so, and not so close as to bring the coefficient of the interaction term into suspicion. Dropping the Gender-Years of Post-secondary Education interaction term from the baseline model has an interesting effect on the Gender coefficient. It decreases by a quarter in absolute value from -.41 to -.30. This change shows that if one does not include a term into the model to take into account the greater ability of women to earn a return on their years of education, whether because of higher grades, better study habits, better motivation, or some peculiarity of the occupational structure women face, one will underestimate the extent of direct discrimination against women.

Do the choices young women and men make of fields of study in college explain any of the discrimination against women because of their gender or the surprising fact that women have a higher rate of return to a year of post-secondary education? A glance at Table 2 shows that men and women do tend to take different fields of study. Men are concentrated in natural sciences and technical fields and business, women in humanities and education. Social sciences are about evenly split between women and men. The data source identifies fields of study of last degree received. Thus, in Table 2 one person may appear as having more than one field of study. If no degree was received at the time of an observation, no field of study is identified. It is known that majoring in natural science and technical fields and business yields greater returns in later earnings than majoring in other fields, so it is reasonable to ask whether any of the discrimination against women is really due to the fact that the subjects women take in college have a lower rate of return than those of men. Column 2 of Table 3 answers this question.

Column 2 of Table 3 contains the results of regressing the logarithm of hourly earnings in 1975 dollars on the baseline model and binary variables for Field of Study and additional binary variables for the interaction between Gender and Field of Study. A glance at Column 2 of Table 3 shows that two fields of study result in earnings significantly different from those of people in the contrast category, people with degrees in fields other than those listed, and people who were once enrolled in post-secondary education but who never received a degree. These are 'natural science, engineering, and technical fields' and 'business.' People majoring in these fields have higher earnings than the others. Two of the interaction terms between Gender and Field of Study are statistically significant. These indicate that women who have taken a degree in the humanities or education have higher earnings, net of the control variables, than men who have taken a degree in these fields. The other interaction terms between Gender and Field of Study are not statistically significant, indicating no statistically significant difference between the earnings of women and men who have taken a degree in these fields, net of the control variables.

What does controlling for Field of Study do to the effect of Gender on Earnings? Does it decrease or eliminate it? Controlling for major field decreases the absolute value of the coefficient of Gender slightly, indicating that some small part of the total negative effect on earnings of being a women is due to Field of Study. The regression coefficient of Gender decreases in absolute value from -.4083 in the baseline model to -.3667 in the regression equation with the Field of Study and the interactions between Gender and Field of Study controlled for. The regression coefficient of -.3667 means that controlling for these additional variables and the variables of the baseline model, women earn 69.3% as much as men do by the hour. We have just explained only three percentage points of the 33 percentage point spread between women's earnings as a fraction of men's earnings, as estimated from the baseline model, and 1.0, the ratio one would expect if gender made no difference in earnings. It is quite clear that most of the negative effect on earnings of being a women has nothing to do with a person's field of study.

Conclusions

This paper assumed at the outset on the basis of the literature that women with at least some college earned less than men with at least some college, because of gender discrimination, and that women had a lower rate of return to a year of post-secondary education than men, again because of gender discrimination. The question was then raised as to whether the choices that women and men make of subjects to study in college account for these gaps in some degree. It turns out that one of the premises of this paper is not correct. Women do not have a lower rate of return to a year of post-secondary education than men. In fact, women's rate of return is higher. It should be noted, however, that this effect is small relative to the much larger negative effect of being female on earnings. Of course, college educated women who are not in the labor force do not enter this study. Were they included with their zero earnings, this finding of women's greater return to a year of education would be surely reversed.

Do the choices women and men make in college affect the gap between their earnings? They do, but the effect is slight. Controlling for major field of study decreases slightly the negative impact on earnings of being female. Controlling for major field of study does not alter in any important way the fact that being female, net of relevant control variables, results in a very substantial reduction in earnings. What do these findings indicate about the nature of discrimination on the basis of gender? They show that the returns to effort expended in investing in learning are roughly the same for both young men and women, and that a substantial part of the differences in earnings between women and men comes from the fact of a behavioral response to gender, namely discrimination against women.

Table I. Distribution of Observations

Entering Analysis by Wave of Study[a] Raw Frequencies

	Women	Men	Total
1967	0	157	157
1968	52	252	304
1969	103	320	423
1970	158	387	545
1971	222	503	725
1972	275	0	275
1973	349	0	349
1975	518	1,212	1,730
Total	1,677	2,831	4,508

[a]Observations are on young women and men, aged 24 to 33, in the labor force and working at least 30 hours a week, and who were once enrolled in post-secondary education, and who are not missing data on a variable involved in the analysis. Since some individuals were observed in more than one wave, the number of individuals studied is smaller than the number of observations. These observations are on 764 young women, and 1,359 young men.

Source: National Longitudinal Surveys of Labor Market Experience, Cohorts of Young Men and Women (Center for Human Resource Research, 1976).

Table 2. Distribution of Observations over Fields of Study.[a]

Field of Study of Most Recent Post-Secondary Degree	Women	Men
Humanities	9.4	0.2
Education	22.6	0.8
Natural Sciences (including mathematics, engineering and technical fields)	9.1	73.1
Business	3.0	12.2
Social Sciences	8.8	10.0
Other fields and people once enrolled in college who have yet to receive a degree[b]	47.2	3.7
Total	100.1% (1,677)	100.% (2,831)

[a]See footnotes of Table 1.

[b]The 'other fields' category is slightly more inclusive for women than men. For both genders, fields of study atypical of the gender are likely to be coded 'other' by the N.L.S. For example, law for women is coded 'other' by the N.L.S. as is 'home economics' for men. Here both are in 'other fields.'

Source: National Longitudinal Surveys of Labor Market Experience, Cohorts of Young Women and Men (Center for Human Resource Research, 1976).

Table 3. Least Squares Regressions: Dependent Variable is the Natural Logarithm of Hourly Earnings in 1975 Dollars.[a]

	Baseline Model	Model with Terms for Field of Study and the Interaction between Gender and Field of Study
Age (in years)	.0434*	.0423*
Years of Enrollment at College Level	.0399*	.0425*
Gender (1=female; 0=male)	-.4083*	-.3667*
Color (1=black; 0=other)	-.0662	-.0593
Prestige of Father's Occupation[b]	.0006	.0004
Associate Degree (1=has degree; 0=not)	.0419	.0135
Bachelor's Degree (1=has degree; 0=not)	.0707	.0546
Master's or Doctorate (1=has such a degree; 0=not)	.0153	.0048
Interaction between Gender and Years of College Enrollment	.0363*	.0359*
Interaction between Gender and Associate's Degree	.0398	.0071
Interaction between Gender and Bachelor's Degree	.0280	.0173
Interaction between Gender and Master's or Doctorate	-.0667	-.1000
Year of Interview[c]	-.0124*	-.0125*

Table 3. continued

Field of Study of Most Recent College Degree		
Humanities	--	-.0384
Education	--	-.0247
Natural Science, Engineering, and Technical Fields	--	.1476*
Business	--	.1172*
Social Science	--	.0232
Contrast Category (other fields and people once enrolled in college who have not yet received a college degree)	--	0.0
Interaction between Gender and Field of Study of Most Recent College Degree		
Humanities	--	.1565*
Education	--	.1170*
Natural Science, Engineering, and Technical Fields	--	.1020
Business	--	.0468
Social Science	--	.0866
Contrast Category (other fields and people once enrolled in college who have not yet received a college degree)	--	0.0
Intercept	.1662	.1942
R^2	.1996	.2211
N^d	2,123	2,123

Table 3. continued

---

[a]Observations are on young women and men, aged 24 to 33, in the labor force and working at least 30 hours a week, and who were once enrolled in post-secondary education; and who are not missing data on a variable involved in the analysis. Dollar values were adjusted to 1975 price levels with the implicit price deflators for "personal consumption expenditures" from the Economic Report of the President (President of the United States, 1977:B-3). Coefficients marked by an asterisk are statistically significant at the .05 level according to an F-test.

[b]Prestige of father's occupation or head of household's occupation when respondent was 14 years of age was measured by the Duncan Socioeconomic Index (Duncan, 1961).

[c]Year of Interview is coded 1 for 1967, 2 for 1968, and so on up to 9 for 1975.

[d]All observations are weighted by the reciprocal of the number of observations on an individual appearing in the sample, making the effective N the number of cases rather than the number of observations in the analysis. Observations are also weighted by the inverse of the probability of an individual's being sampled, i.e., over-sampled strata are constrained to be proportionately a smaller fraction of cases entering an analysis and vice versa for under-sampled cases. The formula used to compute the weight associated with any observation is:

$$\left[(W_{ij})(N)\right] \Big/ \left[\left[\sum_{i=1}^{N}\sum_{j=1}^{J}(W_{ij}/J)\right](J)\right]$$

Table 1 (footnotes) -- continued

where, w_{ij} = sample weight of person i appearing with the jth observation on case i;

N = total number of people in an analysis;

J = total number of observations at different times on person i.

This weighting procedure permits the use of statistics developed for simple random sampling at a single point in time, while using all available information.

Coefficients of regression equations are estimated with the Statistical Package for the Social Sciences (S.P.S.S.) (Nie et al., 1975) so the weights are put in a form usable by S.P.S.S. The equivalent algebraic transformation of the regression equations would be to multiply each equation through by the square root of the weight and then estimate the coefficients with ordinary least squares.

CHAPTER FOUR:

INFERRING WORK EXPERIENCE FROM A LONGITUDINAL SURVEY

ABSTRACT OF CHAPTER FOUR

Earnings determination equations have used "Age minus Schooling minus Five(Six)" as a measure of work experience. This measure has serious problems: 1) it makes many inaccurate assumptions; 2) it prevents simultaneous estimation of age, schooling, and experience effects on earnings; and 3) it ignores longitudinal information if available. This paper develops a measure of work experience from information in a longitudinal survey on whether people are working at the time of the interviews. This new measure is preferable to "Age minus Schooling minus Five(Six) in these three respects.

INTRODUCTION

Chapters 2 and 3 have shown that a person's number of years of school completed, although only the crudest indicator of what the person has learned, is quite adequate to assess the impact of that person's education on his or her occupational achievement. It is apparent that individual variation in in-school learning among those completing a given number of years of schooling is not closely related to success in the labor market. People are apparently like canned goods, judged by their labels, that is their educational credentials, their intellectual contents are largely inaccessible. Now the question is raised of whether learning on-the-job has any effect on earnings and later occupational mobility. In order to answer this question, though, an important methodological innovation has to be made, the inference of the length of time a person has worked from information on whether the person is working at the time of the interviews of a longitudinal survey such as the National Longitudinal Surveys. This chapter is devoted to explaining the theory of how this inference is made. The next chapter will explain the application of the theory.

Human Capital theory posits that on-the-job learning ought to be closely related to marginal productivity and consequently to wage rates. On-the-job learning is usually operationalized as the length of time a person has spent working, a variable called "work experience" or simply "experience" in the literature (cf. Mincer 1974). This chapter examines the concept of work experience and introduces a measure of work experience inferred by interpolation from the interviews of a longitudinal survey. This measure is compared with a cross-sectional estimator of work experience in testing hypotheses about the effect of on-the-job learning on wages.

INDICATORS OF ON-THE-JOB LEARNING

Taking the length of time a person has spent working as the measure of what and how much he or she has learned by working is clearly likely to incur error and ignore many relevant distinctions. Its use is justified only by the absence of a better measure. However crude a time-on-the-job measure of learning, or work experience, may seem, its crudities have ample precedence in the very large number of wage determination studies which take number of years of school completed as the measure of a person's education. A time-on-the-job measure of on-the-job learning can be rather easily modified to take into account the finding of learning theory that learning in a novel situation is more rapid in the earlier periods of exposure than in later periods. If length of time at work is an adequate measure of on-the-job learning, then the square of the term ought to be an adequate measure of any modest and simple departure from linearity in the relationship between learning and the length of time a person works.

Human Capital Theory, taking learning theory into account, predicts a negative coefficient for work-experience-squared, net of work experience.

Age Minus Schooling Minus Five(Six) Years

The most widely used measure of work experience is an estimate of the length of time a person has been out of full-time formal education, and hence an indicator of potential full-time work experience. Its formula is its name, "age minus schooling minus five (or six) years," or A-S-5(6) for short. A person's current age is taken and the number of years he or she completed of formal schooling (or more commonly the number of years of schooling represented by his or her highest grade completed, under the assumption of no skipping or repeating of grades) along with the number of years before the beginning of formal schooling is subtracted. A-S-5(6) has the advantage that it requires only that a person's current age and highest grade completed be known. Since both these items are available from most social surveys and from samples of records prepared by the U. S. Bureau of the Census, A-S-5(6) can be widely applied. However, A-S-5(6) makes the simultaneous estimation of age, schooling, and experience effects on earnings impossible, since A-S-5(6) is simply a linear combination of the other two variables. In the economic literature it is conventional to discuss age or experience effects but not both simultaneously. It is quite conceivable however that age and experience might have different impacts on earnings. This distinction is likely to be of particular importance in the study of the labor force participation of older persons. A-S-5(6) also clearly makes many inaccurate assumptions.

A-S-5(6) assumes 1) that the transition from school to work occurs once and that these activities do not overlap, 2) that people complete a grade in school in the same length of time, 3) that people work the same number of hours a week, the same number of weeks a year, and least accurately, 4) that people have spent the same proportion of time since the end of formal education working. This latter assumption leads to over-estimates of many women's work experience. Older people are going to have much more variance in their working experience than younger people, since those who work regularly and long hours are going to be piling up work experience while those who do not work remain with no work experience. However, the variance of A-S-5(6) within a birth cohort decreases as the cohort ages, since the differences in the time they spent in school are decreasing as a fraction of their lifetimes. A-S-5(6) has the further problem of creating an artifactual correlation between itself as a measure of experience and age and education, two variables from which it is conceptually distinct. Thus, although for teenagers both education and work experience may be increasing with age, teenaged work experience inferred from A-S-5(6) is constrained to be negatively correlated with education. The most important reason for not using A-S-5(6), if longitudinal data are available, is that its use constitutes a massive discarding of longitudinal information on people's work experience. The use of A-S-5(6) can only be justified by the lack of a better measure.

Retrospective Information on Work Experience

Personal work histories can be reconstructed from questions at a single point in time. These data may have problems, however. Memory is fallible and its fallibilities may be correlated with explanatory variables, such as age and education. Older workers may be less able than younger workers to recall their work history accurately since there is more of it and some of it may be quite old and more liable to be forgotten. Also, it may well be that better-educated workers are more able to recall their work histories than the less well educated, as might people who stay in a job for a long time, people with "orderly" careers, that is, a steady increase in wages, and people who work for large organizations which are more likely to keep records and make them available than other employers. Since age, education, length of work experience, tendency to stay with one job or employer, to have an "orderly" career or to work for a large organization are going to be explanations or related to explanations of wages, their relationship to error in the recall of work experience means heteroskedasticity in the wage determination equation. There is also a question whether the amount of error incurred by a retrospective question is tolerable, quite apart from its patterning.

Experience Inferred from a Longitudinal Survey

Griliches (1976) appears to have inferred work experience from a longitudinal survey of labor market experience. The details of the inference are not given. The present paper proposes a measure of work experience based on linear interpolation from what people are doing at the times of the interviews of a longitudinal survey which has little information about work experience between interviews. Work experience is inferred according to the following rules:

1) if person i is working during waves t and t-1, the intervening time, $X_{it} - X_{i(t-1)}$, is added to the cumulator of work experience at time t, X_{iet};
2) if person i is working during one of the waves but not the other, one-half of the intervening time, $X_{it} - X_{i(t-1)}$, is added to the

cumulator of working experience, X_{iet}', and one-half to the cumulator of time out of the labor force, X_{iot}';

3) if person i is not working at either wave, then the whole of the intervening period, $X_{it} - X_{i(t-1)}$, is added to the cumulator of time out of the labor force, X_{iot}.

Linear interpolation is expected to incur error. For example, people working at waves t and t-1 may not have been working in the meantime. Yet, it is inferred that they spent the whole period working. Fortunately, a few things are known about this error in inferring experience. First, it cannot be larger than the period between interviews, $X_t - X_{t-1}$. Secondly, an over- or under-estimate in work experience or time out of the labor force means that time in the other category has been under- or over-estimated by an identical amount. Thirdly, the absolute value of the error incurred by inferring that a person has worked or not worked is determined by the tendency of the person to remain either working or not working during the period between interviews.

The absolute value of the maximum error in the inference of experience between waves t and t-1, $|D_{it}|$, is:

$$|D_{it}| = (X_{it} - X_{i(t-1)}) - (U_{i(t-1)} + U_{it}) \quad (1.1)$$

where,

X_{it} = time from first interview of person i in the longitudinal survey to interview in wave t,

U_{it} = the length of time person i spent in the category he or she was observed in at wave t during the interval between waves t-1 and t,

and,

$$0 < (U_{i(t-1)} + U_{it}) \leq (X_{it} - X_{i(t-1)})$$

so,

$$0 \leq |D_{it}| \leq (X_{it} - X_{i(t-1)}).$$

It can be readily seen that 1) the smaller the interval between observations, and 2) the longer a person remains working or out of work, the smaller the maximum error which can be incurred by inferring experience between the interviews will be. The actual errors in inferring work experience, d_{ie}, and time out of the labor force, d_{io}, are equal and opposite: $d_{ie} = -d_{io}$.

The sums of the d_{ie}'s and the d_{io}'s for a person, $\sum_{t=1}^{t} d_{iet}$ and $\sum_{t=1}^{t} d_{iot}$, are expected to go to zero as the number of waves of the longitudinal survey, t, increases and over- and under-estimates in the inference of work experience or time out of the labor force cancel themselves out. However, some error is to be expected in the cumulator of work experience during the early waves of the longitudinal survey. If work experience is used as an explanatory variable in a regression procedure, the disturbance term will be heteroskedastic (Kmenta, 1971:316), as a glance at equations 1.2, 1.3, and 1.4 shows:

$$Y_t = b_0 + b_1X_{1t} + \ldots + b_{et}X_{et} + e_t \tag{1.2}$$

$$= b_0 + b_1X_{1t} + \ldots + b_{et}(X^*_{et} + \sum_{t=1}^{t} d_{et}) + e_t \tag{1.3}$$

$$= b_0 + b_1X_{1t} + \ldots + b_{et}X^*_{et} + e^*_t \tag{1.4}$$

where,

Y_t = natural log of the hourly wage at wave t

X_{et} = inferred work experience

X^*_{et} = true work experience

$\sum_{t=1}^{t} d_{et}$ = sum of differences between inferred and true work experience from wave 1 to t

$$e^*_t = e_t + b_{et}\left[\sum_{t=1}^{t} d_{et}\right].$$

X^*_{et} and e^*_t are correlated.

The problem with heteroskedasticity can be expected to disappear as the number of waves of the study increases, since $\sum_{t=1}^{t} d_{et}$ is expected to go to zero as t becomes larger. However, the number of waves in a survey of individual labor market experiences has yet to approach any semblance of infinity, so inferred experience during the early waves cannot be simply discarded. The dividing line between an early and a later wave is not clear. Error terms for the inferred experience of different individuals will vary according to their propensity to remain either at work or out of work. The early estimates of inferred experience for some individuals may be quite good, whereas other individuals may require many waves before their work experience may be accurately inferred by linear interpolation.

It is possible to derive a variable which will approximately cancel the magnitude of $\sum_{t=1}^{t} d_{et}$ when divided through equation 1.3 for the observations from the early waves of a longitudinal survey and which will approximate 1.0 in the later waves, leaving them untransformed. This statistic does not cancel the sign of $\sum_{t=1}^{t} d_{et}$, so one is left with a variable intercept for the early waves, if not heteroskedasticity. Let us derive this variable. The starting point of the derivation is with the sum of errors of inference of experience for the ith case at wave t:

$$\sum_{t=1}^{t} d_{iet} = \sum_{t=1} \pm \left[(X_{it} - X_{i(t-1)}) - (U_{i(t-1)} + U_{it}) \right]$$

$\sum_{t=1}^{t} d_{iet}$ equals a sum of positive or negative quantities since over-estimates are likely in frequency and magnitude as under-estimates. For this reason, we expect $\sum_{t=1}^{t} d_{iet}$ to converge to zero. However, should the U_{it}'s, i.e., the length of time a person remains in an observed category of experience, be small, $\sum_{t=1}^{t} d_{iet}$ may take on some substantial non-zero values. $\hat{D}_{it}$ mimics the expected magnitude of $\sum_{t=1}^{t} d_{iet}$ but not its sign, where $\hat{D}_{it}$ is defined as

$$\hat{D}_{it} = \sum_{t=1}^{t} \left[(X_{it} - X_{i(t-1)}) - V_i + 1.0\right]^{1/t} \qquad (1.6)$$

where,

V_i = an individual's average length of time spent in a category of experience, either at work or not at work, between interviews. V_i can be estimated from an occasional wave of the longitudinal survey which ascertains how many weeks a person worked in the previous year, or, if not even that scrap of information about a person's activities between interviews is available, V_i can be estimated as (number of waves during which the person was at work/total number of waves). $(X_{it} - X_{i(t-1)})$,

and,

$$0 < V_i \leq [(X_{it} - X_{i(t-1)})].$$

$\sum_{t=1}^{t} |(X_{it} - X_{i(t-1)}) - V_i|$ can be simplified to $X_{it} - tV_i$, so the statistic $\hat{D}_{it}$ becomes: $(X_{it} - tV_i + 1.0)^{1/t}$. In the first wave, $t = 1$, $\hat{D}_{it}$ is expected to have the same magnitude but not necessarily the same sign as d_{iet}. In immediately following waves it is assumed that it will approach its asymptote at about the same rate as $\sum_{t=1}^{t} d_{iet}$ approaches its asymptote. Division of equation 1.3 by $\hat{D}_{it}$ is expected to approximately cancel the magnitude but not the sign of $\sum_{t=1}^{t} d_{iet}$ in the earliest waves of the survey and to approach 1.0 as $\sum_{t=1}^{t} d_{iet}$ becomes negligible, leaving the equation untransformed.

This procedure for inferring experience and weighting the inferences by what amounts to an estimator of the expected error incurred can be adapted to estimating the values of missing observations. The conventional method of treating cases with missing data in multi-variate analysis is simply to delete them, "listwise deletion." This strategy is self-defeating in a longitudinal survey since one is caught in the dilemma of being more likely to discard information the more one collects. As the number of waves of a longitudinal survey become large, the probability that any one case will at some time be missing data approaches certainty. Listwise deletion, of course, biases a sample and, given the frequency of missing data in most social longitudinal surveys, interferes with the simultaneous use of more than several waves of data.

Let us see how missing data may be interpolated. Suppose a person is missing an observation on whether he or she is at work in one wave of a survey, wave t. One can simply go to the next survey and if it has an observation on whether that person is working, one can infer work experience from wave t-1 to wave t+1 instead of just between wave t-1 and wave t. It can be shown that if there are no missing data after wave t+1, the $\hat{D}_{it}$ statistic for wave t+q is $[X_{i(t+q)} + X_{i(t+1)} - X_{i(t-1)} - (t+q)V_i]^{1/t+q}$. In the example given in this paper, this kind of interpolation is only used to replace missing data on variables necessary to the estimation of work experience. It is potentially applicable to all variables. However, decisions have to be made in estimating $\hat{D}_{it}$ when missing data in one wave must be estimated from two or more other waves and we are not certain of the rules for making these decisions. In the present example, it has been decided to assume that all missing data are being replaced with information from the most distant wave from which any information is taken.

USING WORK EXPERIENCE INFERRED FROM A LONGITUDINAL SURVEY

Let us examine the coefficients of a regression of hourly earnings on inferred work experience and other variables. Since assumptions were made in the derivation of the $\hat{D}_{it}$ statistic, it is of interest to see whether $\hat{D}_{it}$ might be discarded. We will also estimate the impact of experience on earnings using the A-S-6 estimator of experience inferred from longitudinal information. Data are taken from the National Longitudinal Survey of the Labor Market Experience of Young Men (N.L.S.) (cf. Center for Human Resource Research 1976). Only young men 14 to 17 in 1966, the first wave of the survey, are selected. These cases are selected because they can be reasonably assumed to have had no significant work experience before 1966. The technique for eliminating heteroskedasticity introduced by the inference of experience developed here assumes that people have had no work experience at the first observation of the longitudinal survey. However, if it can be shown that division of the wage determination equation by $\hat{D}_{it}$ does not palpably alter the estimates of the regression coefficients, then the practice of using A-S-5(6) to estimate work experience as of the first wave can be justified, and the inference of work experience from longitudinal surveys can be extended to cases likely to have had some work experience as of the first wave.

The 14 to 17 year olds of 1966 represent 2,653 of the 5,225 cases of the N.L.S. survey of young men. Observations on cases are available in eight waves, those of 1966, 1967, 1968, 1969, 1970, 1971, 1973, and 1975. However, the 1973 wave did not ascertain the hourly wage rate. Rather than attempt the wholesale interpolation of wage rates for this wave, it is deleted. However, information on whether a young man was at work in 1973 is used in inferring work experience in 1975. All but the 1970 and 1971 waves of the survey collected information on how many weeks in the previous calendar year a young man worked. All waves have questions on how many hours a week the young man typically worked. Consequently, more information is available than simply whether the young man was at work during the interview, the kind of data set for which $\hat{D}_{it}$ was developed. However, hours worked per week have to be interpolated between interviews. Consequenty, $\hat{D}_{it}$ is necessary to cancel heteroskedastic error from this interpolation, from the interpolation of weeks or hours worked past missing observations on these variables, and from the interpolation of weeks worked per year over longer periods than a year. Work experience is computed as:

$$(X_{it} - X_{i(t-1)})\ (W/52)\ (H/40)$$

where,

W = weeks worked in previous year

H = hours worked in typical week.

An observation on a young man enters the analysis if he is known to be in the labor force at the time of an interview. Observations, not cases, are the unit of analysis. However, observations over time on a case are not independent. The "effective N" (Kish 1967, p. 162) is conservatively taken as the number of cases entering an analysis, rather than the number of observations. Cases are weighted by the reciprocal of the number of observations on them. Cases are also weighted by the inverse of the sample weights, that is, cases from over-sampled strata are constrained to be a proportionately smaller fraction of the cases entering the analysis and vice versa for cases for under-sampled strata. The formula used is given in the footnotes to Table 1. All available observations are used; the basis for hypothesis testing is the number of cases in the sample; the under-sampling or over-sampling of various strata of the universe are compensated for. Regressions are estimated with Statistical Package for the Social Sciences (S.P.S.S.) (Nie et al. 1975) so the weights are put in a form usable by S.P.S.S. The equivalent algebraic transformation of the regression equation would be to multiply each equation through by the square root of the weight and then estimate coefficients with least squares.

The effect of work experience inferred from longitudinal data on the natural logarithm of hourly wage is estimated for equation 2.1. The effect of work experience using the A-S-6 formula on the natural logarithm of hourly wage is estimated for equation 2.2.

$$\frac{Y_t}{D_t} = \frac{b_{30} + b_{31}X_{1t} + b_{32}X_{2t} + b_{33}X_3 + b_{34}X_{4t} + e_{3t}}{D_t} \qquad (2.1)$$

$$Y_t = b_{40} + b_{42}X_{2t} + b_{43}X_3 + b_{45}X_{5t} + e_{4t} \qquad (2.2)$$

where:

Y_t = natural log of hourly earnings at wave t

X_{1t} = age in years at wave t

X_{2t} = education in years of schooling at wave t

X_3 = color (1=black, 0=other)

X_{4t} = work experience inferred from longitudinal data, in full-time equivalent months

X_{5t} = work experience estimated with Age-Schooling-6 formula, in months (work experience before age 14 set to zero).

The coefficients of equations 2.1 and 2.2 are estimated in two steps. The first step involves the estimation of a $\hat{\rho}$, a coefficient of first-order autocorrelation, for each equation. 7,987 observations on 2,420 cases are involved at this step. The second step is the Orcutt transformation using ρ and the estimation of the transformed equations. Since the Orcutt transformation requires the absence of missing data, the number of observations and cases are cut down to 5,340 observations and 1,900 cases. The earnings and education variables are missing data.

The regression coefficients are displayed in Table 1. As can be readily seen, one would arrive at very different conclusions about the impact of age, education, and experience on earnings depending on which measure of experience is used. First, one has to choose between age and education as an explanatory variable if one wants to use A-S-6. One does not with experience inferred from a longitudinal survey. The A-S-6 equation over-estimates the impact of education and experience on earnings because it is not controlling for the effect of age on earnings.

(Table 1 about here)

The computation of $\hat{D}_{it}$ relies on a number of assumptions. It is cumbersome. If it can be eliminated without serious impact on the estimation of regression coefficients, use of work experience inferred by interpolation from longitudinal data would be facilitated. There are two other reasons for wishing to discard $\hat{D}_{it}$. First, its formulation presumes that work experience is known at the first wave of the longitudinal survey. Without the use of A-S-5(6) to estimate work experience in this wave, cases have to be restricted to those which can reasonably be assumed to have no work experience at the first wave. If $\hat{D}_{it}$ can be discarded, then A-S-5(6) estimates of work experience at the first wave can be used and a much larger number of cases can be incorporated into an analysis. Secondly, it is reasonable to ask if work-experience-squared is related to ln hourly wage. Theory suggests it should be, and, net of work experience, should have a negative sign, indicating early work experience having more of an effect on hourly earnings than later work experience. However, division by $\hat{D}_{it}$ of equation 2.1 with a work-experience-squared term does not eliminate the heteroskedastic component of the disturbance term. If division of equation 2.1 by $\hat{D}_{it}$ is superfluous, then a work-experience-squared term can be introduced. Column 1 of Table 2 presents the coefficients of equation 2.1 estimated without division of the equation by $\hat{D}_{it}$. The coefficients of the unweighted equation are virtually identical to those of the weighted equation. While this comparison does not prove that the $\hat{D}_{it}$'s can be discarded with every data set, it strengthens the case for doing so.

(Table 2 about here)

Would it be possible to discard the first several waves of the study, but retain information from them on work experience, and then estimate equation 2.1 without division by $\hat{D}_{it}$? This would be an acceptable procedure if the $\hat{D}_{it}$'s approach 1.0 quickly with successive waves of the study. A glance at Table 3, however, shows that because all information on weeks worked in 1970 and 1971 must be interpolated, the $\hat{D}_{it}$'s do not settle down very quickly to 1.0. Discarding the $\hat{D}_{it}$'s is something of a risk since it cannot be justified on the grounds that the $\hat{D}_{it}$'s have closely approximated 1.0 at any point from the 1966 to the 1975 wave of the study.

(Table 3 about here)

However, since the coefficients of the unadjusted regression equation are virtually identical to those of the adjusted equation, let us make bold enough to see what happens if the $\hat{D}_{it}$'s are discarded and an interpolated work-experience-squared term is added. The results are in columns 2 and 3 of Table 2. There is a strong non-linearity in the relationship between inferred work experience and the natural log of hourly wages. The coefficient of the work-experience-squared term is negative and statistically significant, as expected under Human Capital Theory (cf. Mincer 1974). Its standardized coefficient shows that the non-linearity in the relationship between length of inferred work experience and ln hourly wage is large relative to the relationships between the other variables and ln hourly wage. The coefficient of inferred work experience becomes larger with inferred work-experience-squared in the equation. Indeed a glance at the standardized coefficients in column 3 of Table 2 shows that inferred work experience and inferred work-experience-squared are the largest factors in that regression, larger in their relationship with ln hourly wage than age, education, or being black.

CONCLUSIONS

This paper has reviewed problems with conventional measures of work experience, "age minus schooling minus five (or six) years," or a single retrospective question on job history. A new measure of work experience is proposed which infers work experience from whether a person is working during the interviews of a longitudinal survey. Linear interpolation is used to make this inference. Possible error involved in this inference is discussed and an estimator of it is developed, $\hat{D}_{it}$. Data from the National Longitudinal Surveys of the Labor Market Experience of Young Men are used to compare estimates of the relationship between work experience inferred by interpolation and the natural logarithm of hourly wage and that of work experience from the A-S-6 estimator. The regression equation with work experience inferred by interpolation is divided through by the estimator of error incurred by this process, $\hat{D}_{it}$, to eliminate heteroskedasticity due to error in an explanatory variable. Both equations are Orcutt transformed because of positive autocorrelation. Use of A-S-6 to estimate experience requires that either age or education be excluded as explanatory variables. This paper excludes age, the usual decision in economic studies. The resulting estimates of the relationship of education and experience to earnings are inflated by failure to control for age.

Estimation of the relationship between work experience inferred through interpolation and log wages is attempted without division of the regression equation by $\hat{D}_{it}$. Coefficient estimates are virtually identical between the two equations. If inferred work-experience-squared is added to this equation, both it and inferred work experience have large, statistically significant relationships with log wages. The sign of the work-experience-squared variable is negative, as predicted by Human Capital Theory. Work experience inferred by interpolation from longitudinal data is a more viable indicator of work experience than the estimator A-S-5(6), which ignores the longitudinal information contained in a longitudinal data set, because it makes many fewer inaccurate assumptions and because it permits controls for age and education when a test for the impact of work experience on earnings is performed.

1. Estimates of Coefficients of Equations 2.1 and 2.2[a] (ln $1975)

	Equation 2.1 Unstandardized	Equation 2.1 Standardized	Equation 2.2 Unstandardized	Equation 2.2 Standardized
Age (in years)	.0522*	.4063*	–	–
Highest Grade Completed	.0172*	.0864*	.0678*	.2096*
Color (1=black, 0=other)	-.1508*	-.0715*	-.1355*	-.0619*
Inferred Work Experience (in full time equivalent months)	.0011*	.0666*	–	–
Work Experience from A-S-6 (months)	–	–	.0048*	.2808*
Constant[b]	-.0292		.1756	
R^2	.2736		.1073	
N of Observations[c]	5,340		5,340	
N of Cases[c]	1,900		1,900	
Coefficient of Autocorrelation	.3928		.4813	

[a]Observations are on young men, aged 14-17 in 1966, in labor force and not missing data in two consecutive waves. Estimates are from Orcutt transformed equations.

[b]This quantity equals $b_0^*/(1-\hat{p})$ where b_0^* is the intercept of the Orcutt transformed regression and $\hat{p}$ is the coefficient of first-order autocorrelation estimated over all t-1, t pairs of observations.

Table 1 -- continued

[c]All observations are weighted by the reciprocal of the number of observations on an individual appearing in the sample, making the effective N the number of cases rather than the number of observations in the analysis. Observations are also weighted by the inverse of the probability of an individual's being sampled, i.e., over-sampled strata are constrained to be proportionately a smaller fraction of cases entering the analysis and vice versa for under-sampled cases.

The formula used to compute the weight association with any observation is:

$$[(w_{ij})\ (N)]\ /\left[[\sum_{i=1}^{N}\ \sum_{j=1}^{J}\ (w_{ij}/J)]\ \ (J)\right]$$

where w_{ij} = sample weight of person i appearing with the jth observation on case i;

N = total number of people in an analysis;

J = total number of observations at different times on person i.

This weighting procedure permits the use of statistics developed for simple random sampling at a single point in time, while using all available information. In the Orcutt transformed equation, the weight of the observation in the tth wave rather than the t-1st wave is used.

*Coefficient is statistically significant at the .05 level according to an F-test.

-Not estimated.

Source: National Longitudinal Survey of the Labor Market Experience of Young Men (Center for Human Resource Research, 1976).

2. Estimates of Coefficients of Modifications of Equation 2.1[a] (in $1975)

	Equation 2.1 not divided through by D	Equation 2.1 not divided through by D, and with addition of inferred-work-experience-squared as explanatory variable	
	unstandardized	unstandardized	standardized
Age (in years)	.0493*	.0400*	.2070*
Highest Grade Completed	.0193*	.0193*	.0604*
Color (1=black, 0=other)	-.1290*	-.1232*	-.0575*
Inferred Work Experience (in full-time equivalent months)	.0011*	.0065*	.4031*
Inferred-Work-Experience-Squared	-	-.000037*	-.3038*
Constant	-.0044	.0612	
R^2	.12197	.13237	
N of observations	5,340	5,340	
N of cases	1,900	1,900	
Coefficient of Auto-correlation	.4775	.4693	

[a]See footnotes of Table 1.

3. Means and Ranges of $\hat{D}_{it}$'s by Wave of Longitudinal Survey[a]

	Mean	Minimum	Maximum
1967 (2nd wave)	1.76	1.00	3.90
1968 (3rd wave)	1.66	1.00	2.76
1969 (4th wave)	1.67	1.00	2.36
1970 (5th wave)	2.17	1.63	2.60
1971 (6th wave)	2.09	1.68	2.42
1975 (8th wave)	1.78	1.66	1.98

[a] D_{it} for 1966 is defined as 1.0. However, because of the Orcutt transformation it is not used.

CHAPTER FIVE:

RETURNS TO WOMEN AND MEN FOR WORK EXPERIENCE

ABSTRACT OF CHAPTER FIVE

If work experience is measured by the formula, age-schooling-six(five), it appears that returns in hourly wages to work experience are much larger for men than women. This paper infers work experience from whether people are working when they are interviewed in a longitudinal study, i.e., independently of age and schooling. Women and men aged 24 to 31 and 33 respectively in the U.S. labor force have about the same rate of return to a full-time equivalent month of work experience but men are paid more as they age, regardless of work experience, and women are not. This tendency to pay young men more with age explains all the gap between the hourly wages of women and men. A person's returns to work experience have little relationship to what that person has learned by working, that is, individual variation in job-learned skills has almost no relationship to individual variation in hourly wage.

INTRODUCTION

Sawhill (1973) and King (1977) have found that much of the wage gap between women and men is due to different rates of return to work experience, measured as the length of time a person has worked. There are five principal explanations for why men's work experience results in higher wages than women's. One is that women's qualifications, including job-learned skills, are arbitrarily discounted in the labor market. A second explanation is that, on the average, women tend to work in jobs in which there are fewer valuable skills to learn. Another explanation is that what people learn by working becomes obsolete or is forgotten if a person stops working for a period, which women do with greater frequency than men (Women's Bureau, 1969). These first three explanations are consistent with the neo-classical economics of wage determination, often called "human capital theory" (cf. Becker, 1975; Mincer, 1974). A fourth explanation is that the finding of greater returns to men's experience is an artifact of the way work experience is measured. Most studies measure work experience by an estimate of the length of time a person has been out of full-time education, which likely over-estimates women's work experience and under-estimates that of men. A fifth explanation, not at all consistent with human capital theory, is that men are paid more as they age and acquire dependents, that is, on the basis of status and the needs of that status, not because of the on-the-job learning, and that women are not similarly rewarded. Such a pattern of paying men according to age and number of dependents is a publicly recognized procedure in Japan and part of the institution of <u>Nenko</u> (cf. Karsh, 1976). <u>Nenko</u> is an expression of a patriarchalism not widely thought to exist in any but vestigial form in the United States. This paper examines these explanations of the gap between the returns to the work experience of women and men.

Studies which have found gender differential returns to work experience estimate length of work experience by the formula, 'age minus schooling minus six' (or five, if kindergarten is treated as a year of school), or A-S-6(5) for short. This is an estimate of the length of time a person has been out of full-time schooling. It is only an approximation to the length of time a person has worked, known to be more accurate for men than women since women are less likely to spend all the time since leaving school working (Polachek, 1975). Of course, since length of time spent working is only an indicator of what a person has learned, use of A-S-6(5) is a rough approximation to an indicator. This paper employs a more accurate estimator of the length of time a person has worked, and then refines this indicator of on-the-job learning further by specifying the types of tasks and levels of task complexity a person performs in working, differentiating between people whose work offers many rich and on-going opportunities to learn and those whose work offers little. We will see whether there are gender differential rates of returns to work experience and on-the-job learning measured these ways.

Let us review some of the shortcomings of A-S-6(5) as a measure of work experience. It makes nine unsubstantiated assumptions: 1) that the transition from school to work occurs once and that these activities do not overlap, 2) that people complete a grade of school in the same length of time, i.e., one year, 3) that people work the same number of hours a week, the same number of weeks a year, in 4) jobs which offer exactly the same opportunities to learn the same skills, 5) that people will learn whatever there is to learn by working at the same rate, 6) that people will have spent the same proportion of time since the end of formal education working, 7) that work experience is not distinct from age and education (and so all three variables may not be entered simultaneously as explanatory variables in a regression procedure), 8) that young people cannot acquire work experience if they are engaged in schooling, and 9) that the variance in a birth cohort's work experience decreases as it ages, when, in all probability, it increases.

If longitudinal data are available, use of A-S-6(5) ignores the longitudinal part of the information in the data set. One can improve the measurement of work experience by using the following rules to establish work experience from what a person is doing at the time of an interview and from information on the length of time between interviews:

1. if a person is working in both waves t and t-1, the intervening time is added to his or her tally of work experience,
2. if a person is working at either wave t or wave t-1, one-half of the intervening time is added to the tally of work experience, and,
3. if the person is working during neither wave, then no work experience is added.

Work experience to wave t is the sum of work experience inferred for each period between interviews and up to and including the interview of wave t. Linear interpolation such as by rules #1-3 can incur error, and since this error is an error in an explanatory variable in a regression procedure, it can be shown that this technique of inferring experience introduces heteroskedasticity into the regression. Angle (1979) has worked out a technique for avoiding this heteroskedasticity with weighted least squares and shows that on one longitudinal survey at least that interpolated work experience is more accurate than the A-S-6(5) estimate. This test used information on weeks worked gathered in three consecutive waves of a longitudinal survey to compare with estimates made with the above rules and A-S-6(5). If not being used to test the accuracy of interpolated work experience, information on weeks and hours worked can be incorporated into the interpolated estimates of work experience by adjusting estimates up or down depending on how many weeks a year a person works and how many hours a week. Angle and Wissman (1980) show that as long as there is not a great deal of data missing from a longitudinal survey, estimates of regression coefficients are robust if the special weighting procedure for error incurred through interpolating experience is not used.

It is clear that interpolated work experience is superior to work experience estimated from the A-S-6(5) formula, but is it superior to experience measured by questions on recalled work experience? That test has not yet been made. The answer is clearly contingent on how far back a respondent is asked to remember work experience. Interpolated work experience is based on responses to the question of what the respondent is doing at the time of the interview, involving little effort at recall. Retrospective questions on work experience may ask the respondent for information on work experience many years before. Details blur. A retrospective work experience question can seriously underestimate the experience of those employed in a host of jobs or part-time. In general, recalled work experience may be acceptably well measured for people with stable employment, but for people whose employment has been part-time, haphazard, or who have changed occupations frequently, it may be deficient. Since employment stability and ability to recall past jobs are probably correlated with variables which will be used to explain wages, errors in the measurement of experience will be correlated with other explanatory variables such as age and education in the regression of wages on experience. It is elementary to show that the existence of such errors implies heteroskedasticity.

Improvement in the measurement of work experience lies in the more precise measurement of the length of time a person has worked. Improvement in the measurement of what a person learns by working could take a number of forms. The form offered here is to give information on the type and level of complexity of the tasks which a person performs in working. Use of work experience to explain wages is predicated on the assumption, in human capital theory, that the longer a person works, the more she or he learns relevant to creating wealth in a unit of time, marginal productivity. Some jobs clearly offer more opportunities to learn than others. Jobs with complex tasks obviously offer more opportunities to learn than jobs with simple tasks, which are soon mastered and performed repetitiously with no opportunity for further learning. This indicator of on-the-job learning takes the conventional assumption underlying the use of length of time worked, or work experience, to explain a person's wage rate, i.e., that people learn valuable skills by working, and makes it more accurate, by partitioning time spent in simple jobs where there is little to learn from more complex jobs where there is more to learn. This indicator is an improvement on the conventional indicator, not a radical departure.

The Dictionary of Occupational Titles, (D.O.T.) Third and Fourth Editions (U.S. Department of Labor, 1965, 1977), uses three scales of the complexity of tasks involved in an occupation. These are the "worker function scales," better known as the 'people, data, things' scales. Each measures the highest level of task complexity encountered by a person in working with people, data, or things, three fundamental dimensions of tasks. This paper takes the length of time people work in occupations with different maximum levels of complexity with people, data, and things as the measure of what there is to learn from working in an occupation. Use of this indicator permits the separation of the effect of on-the-job learning on wages from that of other possible effects of longevity of work experience: the effect of seniority either formalized by contract or by the informal practice of rewarding people with more experience or by taking a person's last wage as evidence of their marginal productivity, giving the experienced the option of keeping their present wage or applying for a better paid opportunity and gradually raising their wages.

The worker function scales originally were developed at the end of World War II to facilitate the demobilization of British troops. Individuals were told what civilian occupations their military occupations had given them experience in by matching the profiles of their military occupation with civilian occupations on these scales. The people, data, things scales have 'use validity', that is, they were devised in necessity as a thorough yet simple analysis of an occupation's tasks. They have proven in practice to be a useful way of analyzing what tasks are involved in a job. The scales were adopted by the U.S. Employment Service in the early

1950's with some modification (Broom et al., 1977). These modified scales are intended to order different tasks from the less complex to the more complex and to be hierarchical, that is, for the more complex scale positions to subsume the less complex scale positions. See Figure 1. Question has been raised whether this hierarchical relationship exists, particularly for the people scale (Walther, 1960). Broom et al. (1977) have suggested that the ordinality and hierarchicality of the scales would be less open to question if there were fewer points on them. In their use of the scales, they combined categories. Their method of combining categories has been adapted in this paper. Kohn and Schooler (1973), Temme (1975), as well as Broom et al. (1977) have recognized that the people, data, things scales are the only measure of job content and complexity now extant which can be used in conjunction with national labor force surveys, and, consequently, are uniquely valuable tools for the analysis of occupations despite what imperfections they may have. Temme (1975) aggregated the people, data, things scores of the D.O.T. to the level of the 1960 Census occupation codes, a much grosser occupational classification scheme. His people, data, things scores for each of the 1960 Census occupation codes are averages weighted by the number of people in each of the D.O.T. occupations in the October, 1966 Current Population Survey, conducted by the U.S. Bureau of the Census. The occupations of respondents of this survey were coded by both the Department of Labor coders who used the Dictionary of Occupational Titles codes and the U.S. Bureau of the Census occupation coders who used the 1960 Census codes. As long as the people, data, things scales are ordinal and have seven or more categories, it is acceptable to assign interval numbers to each category (Labovitz, 1970, Kim, 1975).

Data.

The National Longitudinal Surveys of Labor Market Experience (N.L.S.) of young women and young men use the 1960 Census occupation codes and so permit the matching of a D.O.T. occupation code and its data, people, things scores (cf. Center for Human Resource Research, 1976). The surveys of the young women and men are used instead of the older men and women because it is easier to estimate the total time spent working of the young than of the old since there is less of it and one knows that, by Census definition, a person's experience in the civilian labor force on his or her fourteenth birthday is zero. There are 5,225 young men in the N.L.S. study of young men. These were aged 14 to 24 years in 1966. They were interviewed in 1966, 1967, 1968, 1969, 1970, 1971, 1973, and 1975. They were 23 to 33 years of age in 1975. There are 5,159 young women, aged 14 to 24 years as of 1968, and interviewed in 1968, 1969, 1970, 1971, 1972, 1973, and 1975. The young women were 21 to 31 years of age in 1975. Only observations on people who have clearly begun to make the transition from full-time study to full-time work are entered into the analysis of returns to experience. Many economists prefer a "Commencement" model of work activity in which the transition from education to work is assumed to be instantaneous (Mincer, 1974). The "Commencement" model is a poor description of reality, but the need to draw an arbitrary line between those primarily engaged in education and those primarily engaged in work remains since returns to experience can only be properly measured for those who have committed themselves or have been permitted to commit themselves to working. It is useful to have a minimum age for inclusion in the analysis since the relationship between earnings and education or experience may be quite different for the well and poorly educated at a young age from what it will be during most of their working lives. To avoid these problems this study used the following criteria to make the distinction between those primarily engaged in work and those primarily engaged in something else: for inclusion, an observation on a young person must be on one who is at least 24 years of age, and works at least 30 hours a week in his or her current job. If data on whether the person is working 30 hours a week are missing, the number of hours worked a week is estimated by linear interpolation from other waves of the longitudinal survey. Observations on people making less than 25¢ an hour are discarded since the natural logarithm of hourly wage is taken as the dependent variable and this transformation does not yield meaningful values for hourly observations near zero. Observations on people with data missing on a variable which enters the analysis are also discarded.

Observations, not people, are selected for analysis. If a young person does not meet the criteria for inclusion in one wave of the N.L.S., he or she may in another wave. Table 1 displays the numbers of observations on women and men which enter the analysis by each wave of the survey. As can be seen, most of the observations are from the more recent waves since the respondents are older in these and more likely to be 24 years of age or older and working. There are no observations on women in 1966 and 1967 since they were not surveyed in these years, and no observations on men in 1972 for the same reason. There are no observations on men in 1973 since information on their hourly pay rate was not obtained in that wave.

However, information on hours and weeks worked in 1973 is used in estimating work experience. There are 4,106 observations on 1,780 young women and 7,180 observations on 2,890 young men in the analysis.

Observations on the same person over time are not independent. The "effective N" (Kish, 1965:162) of the sample for the purpose of hypothesis testing is conservatively taken to be the number of cases, not the number of observations, in an analysis. Cases are weighted by the reciprocal of the number of observations on them. Cases are also weighted by the inverse of the sample weights, that is, cases from over-sampled strata are constrained to be proportionately smaller fraction of the cases entering the analysis and vice versa for cases from under-sampled strata. See the footnotes of Table 2 for the formulation of this technique. This weighting technique permits the maximum use of available information while permitting the use of statistics which assume simple random sampling in which the number of independent contributions of information is the same as the number of cases observed.

It is possible that there are autocorrelated disturbances in regression equations estimated with this data set. However, computing coefficients of autocorrelation in order to perform an Orcutt transformation would entail the loss of a massive proportion of the data set. Computing a coefficient of autocorrelation requires an estimate of the disturbances in both the 't'th and 't-1'st waves. Thus all observations on young men in 1975 would be lost since the 't-1'st wave, 1973, is missing data on hourly pay rate. Also, computation of the coefficient of autocorrelation at t and t-1 requires that both observations meet the criteria for selection into the analysis and that neither is missing an observation. The conditions decimate the available observations. Whatever gain in efficiency of estimation that would be gained by an Orcutt transformation is offset by the biasing of the data set through the loss of observations.

Analysis

We want to find out whether women and men appear to have different rates of return to their work experience because of the way that experience is measured either by interpolation from longitudinal information or by the A-S-6 formula. We want to examine the relationship between gender and work experience but we want to simultaneously impose a number of controls to eliminate the effect of variables which are very likely more basic in wage determination than the factors which create any difference in the rate of return by gender. These very basic factors we want to control for are of two kinds: 1) measures of skill acquisition, and 2) basic ascribed statuses which people react to and discriminate on. Skill acquisition is measured by three variables: formal education (highest grade completed in school), whether the person has received on-the-job training, and work experience. The basic ascribed statuses which are controlled for are: age, color, parental social class, and gender. The explanation of wages in terms of these factors is called the "baseline model," because it creates a baseline from which the effect of the interaction of gender and work experience on wages is measured. Also included in the baseline model is an experience-squared term, intended to measure the widely observed tendency for work experience to yield a smaller return per month as it becomes longer and longer.

There are two baseline models, one for each way experience is measured. Equation #1 is the baseline model for work experience interpolated from a longitudinal survey. Equation #2 is the baseline model for work experience measured by the A-S-6 formula. When A-S-6 is used as an explanatory variable in a regression equation, age or schooling has to be excluded as an explanatory variable to prevent multicollinearity. The A-S-6 variable use here has been non-linearly transformed, i.e., all experience before the age of 14 is set to zero; so it is technically possible to enter age, schooling and this transformed to A-S-6 variable simultaneously as explanatory variables. However, this A-S-6 contains little information beyond what is in age and schooling and so one of the three is excluded. We follow the convention among economists of pretending that age is irrelevant to wages and exclude age from the regression. The natural logarithm of hourly wage is taken as the dependent variable, instead of hourly wage untransformed, for statistical reasons (Stolzenberg, 1975; Griffin, 1978).

$$\ln Y_t = b_0 + b_1X_{1t} = b_2X_{2t} = b_3X_{3t} + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_{7t}$$
$$+ b_8(X_{7t})^2 + e_1 \qquad \text{eq.1}$$

$$\ln Y_t = b_0 + b_1X_{1t} + b_2X_{2t} + b_4X_4 + b_5X_5 + b_6X_6 + b_9X_{9t}$$
$$+ b_{10}(X_{9t})^2 + e_2 \qquad \text{eq. 2}$$

where,

ln Y = natural logarithm of hourly wage in current job

X_{1t} = highest grade completed

X_{2t} = on-the-job training (1=some training, 0=none)

X_{3t} = age in years

X_4 = color (1=black; 0=other)

X_5 = gender (1=female, 0=male)

X_6 = parental social class (Duncan socio-economic index score of occupation of head of household when respondent was 14).

X_{7t} = length of work experience interpolated from longitudinal data (in full-time equivalent months)

X_{9t} = length of work experience estimated from A-S-6 (in years)

Length of work experience measured by interpolation from people's work activity at the time of their being interviewed in a longitudinal survey includes part-time work and work before they left full-time schooling. It is adjusted by the number of weeks and hours a person works. Each unit in which interpolated experience is measured is the full-time equivalent of a straight-time month, i.e., four weeks of forty hours. Interpolated work experience is correlated only .59 with length of work experience measured by the A-S-6 formula. The average lengths of work experience for the observations on women and men who are in the regression analysis are 67.6 full-time equivalent months for women and 102.9 full-time equivalent months for men when inferred by interpolation from a longitudinal survey and 83.0 months for women and 93.1 months for men when estimated by the A-S-6 formula.

The A-S-6 formula over-estimates women's experience and under-estimates that of men. The use of A-S-6 to estimate the difference between the rates of return to experience of women and men is questionable.

Let us see whether the use of A-S-6 leads to different estimates of the rate of return to experiences from what would be arrived at with measuring work experience by interpolation from a longitudinal survey. Let us compare the unstandardized coefficients of the two variables. The unstandardized coefficients of interpolated experience are in column 1 of Table 2 and the unstandardized coefficients of A-S-6 are measured in years, so the comparison requires making the units of measurement the same. We multiply the coefficient of interpolated experience, .00326, by 12 and the coefficient of interpolated experience squared, -.000011, by 144. Even though the two measures of work experience are only correlated .59, the estimates of returns to experience and the rate of decrease of this return with greater experience are almost the same.

	Interpolated Experience	A-S-6
Experience (in years)	.03912	.04367
Experience Squared (in years squared)	-.00158	-.00138

Interpolated experience estimates the rate of return to a year of work experience to be an $e^{.03912}$ or 4 percent increase in hourly wage whereas the A-S-6 estimate of the increase in wages with an additional year of experience is $e^{.04367}$ or 4.5 percent. However, the equation with A-S-6 does not control for the increase of wages, among young workers, with age, estimated to be 1.3% per year in the equation with interpolated experience.

Estimating work experience by interpolation from longitudinal data shows that, among young workers, age has an effect on wages independent of work experience. With A-S-6 the age effect is lumped together with that of experience. There is, however, no radical discrepancy between the results of the two methods of estimating rate of return to experience. The other coefficients of explanatory variables are about the same between the equations with different ways of estimating work experience, except of course for the coefficient of age, which is arbitrarily set to zero in the A-S-6 equation. If one is just examining returns to experience, the lengthy programming which goes into estimating work experience by interpolation is hardly worth the effort since it yields estimates which are virtually identical to those available with the much simpler, if cruder, A-S-6 estimator.

However, if the interaction term between gender and experience is formed and entered as an explanatory variable, one sees that the technique of estimating experience does affect one's conclusions dramatically. Use of A-S-6 to estimate differential returns to experience by gender results in an estimate of differential returns to experience by gender showing a massive discounting of the experience of women. See columns 7 and 8 of Table 2. Use of interpolated experience leads to a much smaller estimate of the discounting of women's experience. When adjusted for the difference in units, (interpolated experience is in months, the A-S-6 estimate in years), one sees that the estimate of the discounting of women's experience with interpolated experience is only 22 percent of the magnitude of the effect estimated with A-S-6. The A-S-6 equation, however, produces a smaller direct negative effect on wages of being female, .13222, as opposed to .27080 with interpolated experience. Both techniques show costs to women for their gender, but the techniques suggest different ways in which this cost is exacted. The equation with interpolated experience suggests that most of the cost is simply a response to gender not a discounting of skills women acquired by working, and vice versa for the equation with A-S-6.

Is there any way to resolve this apparent indeterminacy? It should be remembered that A-S-6 is really a measure of age after the end of schooling not a direct measure of work experience. Since the variable, age, must be left out of the A-S-6 equation, A-S-6 may, statistically, be acting as a surrogate for age. With interpolated experience, one can separate the effects of experience from those of age, so it is possible to test to see whether there is a massive interaction between gender and age on wages. It is possible that the large coefficient between gender and A-S-6 is reflecting such a large interaction. The results of the addition of an age-gender interaction to the baseline model with interpolated experience is displayed in columns 1 and 2 of Table 3. There is a massive interaction between gender and age. Its coefficient shows that among young workers, men are paid more as they age. This equation also shows that the direct effect of being female on wages is positive, but that this positive effect is overwhelmed by the much larger tendency to pay men more as they age, not because of experience, but because of age <u>per se</u>. Controlling for the effect on wages of this interaction between gender and age hardly alters the estimate of returns to experience. As columns 3 and 4 of Table 3 show there is no statistically significant interaction between gender and work experience net of the interaction between gender and age. The returns to experience are exactly the same for women as men. It is age not experience which is rewarded in men much more than women. As columns 1 and 2 of Table 3 show there is no tendency for the earnings of both sexes to increase with age. The coefficient of age is not statistically significant net of the gender-age interaction. Only young men are paid more because of their increasing age, not young women. This finding is consistent with the existence of a covert Nenko system in the United States in which young men's wages increase as they age and acquire dependents. It is this effect which Sawhill (1973) and King (1977) observed and thought to be the result of a discounting of women's experience. The finding of a massive interaction between gender and age, net of work experience, on wages contradicts the assumption of human capital theory that a person's wages are mostly returns on that person's human capital. We have found that a very important component of a person's wages is determined by social status, not human capital.

Since there is only a small tendency for young men to receive more pay than young women for a month of experience, net of the custom of paying young men more because of their age, the original reason for our looking into returns to on-the-job learning has been removed. We had thought that there might be a possibility that men had acquired experience in more complex levels, on the average, than women, a possibility which could have explained a higher rate of return to their experience. However, we have found something which flatly contradicts the established human capital model of wage determination, a massive tendency to pay young men more than young women as they age, because of their age. We can think of no possible deduction from human capital theory which can explain why young men should develop human capital with age, controlling for amount of work experience; and young women not. Let us look further. Perhaps the human capital theory deduction that anyone is paid more because of their individual level of skill, and thus their individual marginal productivity, is not borne out by data.

Table 4 displays the results of three regressions. Each regression is identical to equation #1, the baseline model of wages, which uses interpolated experience and interpolated experience squared as explanatory variables. The estimates of the coefficients of this regression model are given in columns 1 and 2 of Table 2. The only difference is that the regression equations of Table 4 subdivide interpolated experience into the lengths of time a person has worked at different levels of maximum task complexity. There is one equation for tasks involving people, another for data, and a third for tasks involving things.

If experience with more complex tasks had more of a positive effect on wages than experience with less complex tasks, one would expect an order to the coefficients of work experience at a particular task complexity. The coefficients of work experience at low task complexity ought to be smaller than the coefficients of work experience at high task complexity and the coefficients of experience at low task complexity and high task complexity ought to be different from the coefficient of work experience undifferentiated by the level of task complexity at which it was acquired, or .00326 (column 1, Table 2). There is no clear patterning in the coefficients of work experience at tasks of different complexities in Table 4. None of these coefficients are statistically different from .00326 except the coefficient of experience with people at the highest level of complexity, which is, against prediction, lower than all the other coefficients, and the coefficient of experience with things at the highest level of task complexity. There is very little evidence here that people with more complex skills are paid more than people with less complex skills. There is little or no market response to individual skill differences. These findings do not contradict the finding that people with longer work experience are paid more because of it. It is clear, though, that the effect of experience on wages is unrelated to a process of learning job skills.

CONCLUSION

This chapter began with the intention of investigating whether apparent differences between the rates of return to women and men for work experience were an artifact of the way work experience is measured. On the way to our conclusions we have watched our initial premise disintegrate. We started with the acceptance of human capital theory, the very widespread and conventional interpretation of individual wage differences as the result of individual skill differences. We thought that there might be a good chance that labor markets do not operate in an extremely discriminatory fashion with regards to returns to experience but that the usual indicator of work experience, A-S-6, simply over-estimated the work experience of women and under-estimated the work experience of men, thus showing that women's work experience is apparently discriminated against. We expected to find some discounting of the work experience of women with our measure of work experience interpolated from longitudinal data, but less than that shown by the A-S-6 estimator of experience. We expected to show that men, on the average, work in jobs with more complex tasks than women, generating more valuable experience, and accounting for the gap between the genders in returns to experience. Instead we find there is an enormous positive effect of age, net of work experience, on the wages of young men that does not exist for young women. This effect accounts for all men's tendency to earn more than women by the hour left unexplained by the other variables of the baseline model, equation #1. It appears that there is a covert Nenko system in the U.S. It is the pattern of men's earning more as they age because of their age which produced the large apparent gender-experience interaction on wages, when A-S-6 was used as the measure of work experience. A-S-6 does not permit the separation of age from experience effects. The tendency of young men to receive more money with age because of age is not a deduction from human capital theory and suggests the irrelevance of human capital theory to the reality of wage determination. We have further shown that individual variations in on-the-job learning, as we measure it, bear no relationship to individual variations in wages. We began by accepting human capital theory as a truism; we end by rejecting its relevance. There is no labor market for individual differences in on-the-job learned skills.

It has been found that work experience pays and apparently pays at about the same rate whether it is estimated by interpolation from longitudinal data or by the A-S-6 formula. It is not on-the-job learning which accounts for this effect since the effect does not vary much across work experience at different levels of task complexity and opportunities to learn skills. It may be that a person's length of work experience is a proxy for job-learned skills and that there is a market for differences in this proxy rather than in the skills the proxy is supposed to represent. Or we could be simply measuring the advantages of seniority, whether formalized in a contract or the result of simply having more time to worm one's way into a better deal. Explaining why wages increase with the length of time a person has worked but why they are unrelated to opportunities to learn skills on the job is beyond the scope of this paper. We can only say that the effect on wages of work experience, measured as length of time worked and net of age, operates in almost the same way for women and men. We have found that, though crude, the A-S-6 formula yields approximately accurate estimates of returns to work experience. A-S-6 should not, however, be used to estimate gender differential returns to work experience since it confounds the gender-experience interaction with the large gender-age interaction.

Table 1. Distribution of Observations Entering Analysis by Wave of Study[a]
Raw Frequencies

	Women	Men	Total
1966	0	276	276
1967	0	516	516
1968	147	756	903
1969	279	919	1,198
1970	422	1,005	1,427
1971	582	1,323	1,905
1972	719	0	719
1973	837	0	837
1975	1,120	2,385	3,505
TOTAL	4,106	7,180	11,286

[a]Observations are on young women, aged 24 to 31, and men, aged 24 to 33, working at least 30 hours a week, and who are not missing data on a variable involved in the analysis. Since some individuals were observed in more than one wave, the number of individuals studied is smaller than the number of observations. These observations are on 1,780 young women and 2,890 young men.

Source: National Longitudinal Surveys of Labor Market Experience, Cohorts of Young Men and Women (Center for Human Resource Research, 1976).

Table 2. Least Squares Regressions (Dependent variable is the natural logarithm of hourly wage in current job in 1975 dollars[a]; N= 4,670[b])

	Baseline Model with Interpolated Experience (full-time equivalent months)[c]		Baseline Model with Interpolated Experience and Its Interaction with Gender		Baseline Model with A-S-6[c] (in years)		Baseline Model with A-S-6 and its Interaction with Gender	
	unstandardized	standardized	unstandardized	standardized	unstandardized	standardized	unstandardized	standardized
Highest Grade Completed	.05564*	.30916*	.05621*	.31230*	.07267*	.40375*	.07458*	.41437*
On-the-Job Training (1=some training, 0=none)	.06108*	.06582*	.05999*	.06465*	.06507*	.07012*	.06234*	.06718*
Age (in years)	.01275*	.05284*	.01236*	.05122*	-----	-----	-----	------
Color (1=black, 0=other)	-.14529*	-.09728*	-.14398*	-.09641*	-.14821*	-.09924*	-.14625*	-.09793*
Gender (1=female, 0=male)	-.32239*	-.34101*	-.27080*	-.28645*	-.36741*	-.38864*	-.12500*	-.13222*
Parental Social Class[d]	.00044	.02301	.00045	.02352	.00043	.02265	.00042	.02203
Experience	.00326*	.29150*	.00321*	.28690*	.04367*	.28749*	.02923*	.19243*
Experience Squared	-.000011*	-.19628*	-.000013*	-.23439*	-.00138*	-.13550*	-.00196*	-.19305*
Interaction between Gender and Experience	------	------	-.00069*	-.08445*	-----	------	-.03739*	-.33313*
Constant	-.02969		-.01428		.00776		.10344	
R^2	.30820		.31624		.29969		.31327	

Table 2 continued

[a]Observations are on young women aged 24 to 31 and young men aged 24 to 33, working at least 30 hours a week, and who are not missing data on a variable involved in the analysis. Weeks worked per year and hours per week are interpolated if information is missing (cf. Angle, 1979). Dollar values are adjusted to 1975 price levels with the implicit price deflators for "personal consumption expenditures" from the Economic Report of the President (President of the United States, 1977:B-3). Coefficients marked by an asterisk are statistically significant at the .05 level according to an F-test.

[b]All observations are weighted by the reciprocal of the number of observations on an individual appearing in the sample, making the effective N the number of cases rather than the number of observations in the analysis. Observations are also weighted by the inverse of the probability of an individual's being sampled, i.e., over-sampled strata are constrained to be proportionately a smaller fraction of cases entering an analysis and vice versa for under-sampled cases. The formula used to compute the weight associated with any observation is:

$$\left[(W_{ij})\ (N)\right] \Big/ \left[\left[\sum_{i=1}^{N}\ \sum_{j=1}^{J}\ (W_{ij}/J)\right] \Big/ (J)\right]$$

where,

W_{ij} = sample weight of person i, appearing on the jth observation on person i;

N = total number of people in the analysis;

J = total number of observations in different waves on person i.

This weighting procedure permits the use of statistics developed for simple random sampling at a single point in time, while using all available information in a longitudinal data set.

Coefficients of regression equations are estimated with the Statistical Package for the Social Sciences (S.P.S.S.) (Nie et al., 1975) so the weights are usable by S.P.S.S. The equivalent algebraic transformation of the regression equations would be to multiply each equation through by the square root of the weight and then estimate the coefficients with ordinary least squares.

[c]Work experience before age 14 is defined as zero. Interpolated work experience incorporates an A-S-6 estimate of work experience as of the first wave of the N.L.S. survey.

[d]Prestige of father's occupation or head of household's occupation when respondent was 14 years of age was measured by the Duncan Socio-economic Index (Duncan, 1961).

Table 3. Least Squares Regressions (Dependent variable is the natural logarithm of hourly wages in current job in 1975 dollars[a]; N= 4,670[b])

	Baseline Model with Interpolated Experience and Interaction between Age and Gender		Baseline Model with Interpolated Experience and Interaction between Age and Gender as well as Interaction between Experience and Gender	
	Unstandardized	Standardized	Unstandardized	Standardized
Highest Grade Completed	.05335*	.29644*	.05338*	.29660*
On-the-Job Training (1=some training, 0=none)	.06146*	.06623*	.06142*	.06619*
Age (in years)	-.00707	-.02929	-.00699	-.02896
Color (1=black, 0=other)	-.14524*	-.09725*	-.14520*	-.09722*
Gender (1=female, 0=male)	.48227*	.51014*	.48019*	.50794*
Parental Social Class[d]	.00045	.02364	.00045	.02365
Experience	.00340*	.30386*	.00340*	.30365*
Experience Squared	-.000012*	-.22667*	-.000012*	-.22777*
Interaction between Gender and Experience	-----	-----	-.00002	-.00276
Interaction between Gender and Age	-.03157*	-.87632*	-.03142*	-.87222*
Constant	.50321*		.50121*	
R^2	.31146		.31146	

See footnotes of Table 2.

Table 4. Least Squares Regressions (Dependent variable is the natural logarithm of hourly wage in current job in 1975 dollars[a]; N=4,670[b])

	Interpolated Experience Divided into Lengths of Time Worked at Particular Levels of Maximum Task Complexity with:					
	People		Data		Things	
	Unstandardized	Standardized	Unstandardized	Standardized	Unstandardized	Standardized
Highest Grade Completed	.06587*	.37686*	.04961*	.27567*	.05778*	.33715*
On-the-Job Training (1=some training, 0=none)	.06381*	.06861*	.05037*	.05429*	.05611*	.06033*
Age (in years)	.02444*	.11518*	.01107*	.04585*	.02264*	.10670*
Color (1=black, 0=other)	-.16207*	-.14708*	-.12494*	-.08366*	-.15798*	-.14337*
Gender (1=female, 0=male)	-.30579*	-.31635*	-.33847*	-.35803*	-.30306*	-.31353*
Parental Social Class[c]	.00080	.04015	.00023	.01216	.00085	.04262
Level of Complexity of Work Experience (full-time equivalent months)						
level unknown	.00350*	.26456*	.00326*	.19889*	.00343*	.25913*
1 (low)	.00344*	.26727*	.00171*	.10133*	.00318*	.20835*
2	.00341*	.19218*	.00358*	.21446*	.00289*	.19129*
3 (high)	.00069	.02275	.00407*	.22015*	.00485*	.21823*
Experience Squared	-.000013*	-.30533*	-.000010*	-.19185*	-.000013*	-.30812*
Constant	-.41420*		.10265		-.27925*	
R^2	.34714		.31840		.34614	

See Footnotes of Table 2.

CHAPTER 6:

THE LOOSE CONNECTION BETWEEN LEARNING AND EARNINGS

In June, 1977 the National Institute of Education (NIE) sent out a request for research proposals on, among other topics, how vocational guidance counseling could be improved by taking into account the fact that most people change jobs during their working lives (National Institute of Education, 1977). Traditional vocational guidance counseling attempts to identify what single type of work young people may be interested in, ignoring the possibility that many young people will likely hold a variety of jobs during their working lives. It seemed to me that part of this issue was the question of whether learning on the job, like learning in school, had an impact on the nature of a person's later occupation, in particular, on his or her earnings. If such were the case, then a prudent vocational guidance counselor might want to advise a student to consider the value of occupational experience along with current pay in choosing an occupation. It would be advice to keep on investing in skill acquisition even beyond graduation from school. Whereas it was conventional to view the impact of skill acquisition on a person's work to occur once, after initial graduation from full-time formal education, this project's research proposal viewed the impact of skill acquisition as an on going process. Instead of attempting to advise vocational guidance counselors on what skills were needed for particular advantageous occupational sequences, ones leading upward in terms of pay, prestige, and interest, a task which becomes less feasible the more finely one specifies occupations (because the number of sequences increases geometrically), this project proposed to help advise vocational guidance counselors on what occupations themselves offered opportunities for learning which would have a valuable pay-off later.

The proposal discussed the phenomenon of the experience-earnings curve, the well-known tendency of the earnings of better-educated people to rise more quickly with work experience than those of less well-educated people (Mincer, 1974). The proposal hypothesized that this phenomenon is due to better-educated people managing to enter jobs in which there is more to learn and that the increase in their earnings is a market response to the skills they acquire by working in jobs with more to learn. The purpose of the research was to lay the basis of a computer simulation of young people's entry into the U.S. labor force, a simulation which could be used to identify, from among likely alternative entry occupations, the ones in which a young person would acquire the most valuable experience, the kind of experience which would lead to increased earnings in the future. Essentially this simulation would give a young person an estimate of the value of the experience he or she would be acquiring in different occupations he or she would be entering. This simulation would be a valuable tool for vocational guidance counselors. This project only envisioned the development of a prototype model.

Two essential research questions had to be addressed in order to develop this simulation. First, it had to be found out what there is in formal education which affects occupational achievement, that is, occupational prestige and earnings. Secondly, it had to be found out whether on-the-job learning affects later job characteristics. Answering both these questions entailed solving methodological problems. A way had to be devised to use all available information in a longitudinal survey, and a method of measuring on-the-job learning also had to be devised. These problems were successfully dealt with. Their solutions are described in the foregoing chapters.

Chapter 2 showed for young adults that, as far as it was possible to measure, the principal part of a person's education which affects his or her later occupational achievement is primarily the number of years of schooling a person completes. The subjects a person studies are a distant second in their influence on occupational achievement. It is remarkable given the enormous diversity in what people learn in school and how well they learn it that such a simple measure of a person's education should explain so much of the impact of education on a person's occupational earnings and prestige. Chapter 3 examined the issue of the impact of subject area of study on earnings further. It raised the possibility that, among college-educated people, it might be the decision to major in one field rather than another which explains the gap between the earnings of college-educated women and men, presently employed. It is not. The decision to major in one field or another does explain some of the earnings gap but not much. It appears that people coming out of schools are like canned goods as far as the reaction of the labor market to them is concerned. Their educational credentials are their labels. Those with the same credentials may have all kinds of differences in their knowledge but these differences are inaccessible. Their identity in the labor market is their label. Their individual differences in knowledge do not affect their earnings much.

Chapter 4 introduced the technique of inferring the length of time a person has worked from whether he/she is working during the interviews of a longitudinal survey. This technique is the solution to a methodological problem crucial to this project. This innovation is applied in Chapter 5, an examination of returns to work experience and on-the-job learning and the question of why the earnings of young men tend to apparently rise more quickly than those of young women with experience. Chapter 5 shows that both young women and men are paid more as they accumulate work experience, the total length of time they have worked, albeit with women deriving somewhat less advantage from experience than men. Chapter 5 finds that a large gap opens between the wages of women and men as time goes on but not because the experience of women is discounted, as has been supposed. Men are paid more as they age because of age, not experience, which is controlled for; women are not. This effect explains all the difference between young women's and men's wages on the average among young adults.

Chapter 5 introduces a measure of on-the-job learning distinct from work experience, the conventional indicator to date. The new indicator is simply an extension of the logic of the old one, which was that the longer a person worked the more they learned. The new indicator assumes that working in a more complex job results in more learning than working in a less complex job. Chapter 5 shows that there is no statistically significant difference between the average rate of increase of pay with experience and the rate of increase of pay with experience with complex or simple tasks. Individual variations in on-the-job learning are not responded to by the labor market. However, average rates of learning, represented by simply the length of time a person has worked, may be responded to by the labor market. Perhaps, the increase in wages with experience is simply due to the effect of seniority whether formalized in a contract or simply the informal accumulation of power at the work place, or some other mechanism which accounts for people's wages increasing with experience. The issue cannot be settled with the data at hand.

Since on-the-job learning has no effect on earnings independent of simply the length of time a young person has worked, there is no need for individualized simulations of young people's entry into the labor force. The advice is the same for all with a given level of education, facing entry into the labor force: start work early in as highly paid an occupation as possible and work continuously in it. There is no need to consider the future value of experience independently of current wages. However, it is clear that to the extent some career lines deviate from the average increase in wages with experience, it is a facet of the career line itself, not of the skills of the individuals in that career line. The task facing a young person then is to find entry jobs which lead to better paid positions, not because of learning acquired in the entry job, but because that career line and its increasing wages are part of the social structure of the labor force. A young person should keep in mind that the connection between learning and earnings is a loose one.

REFERENCES

Alwin, Duane F.

1974 "College effects on educational and occupational attainments." American Sociological Review 39:210-223.

Angle, John

1979 "Work and earnings: Cumulative experience method of analysis of longitudinal surveys." Sociological Methods and Research 8:209-231.

Angle, John and David Wissmann

1980 "Inferring work experience from a longitudinal survey." Unpublished manuscript.

Ashenfelter, Orley and Joseph D. Mooney

1968 "Graduate education, ability and earnings." Review of Economics and Statistics 50:78-86.

Astin, Alexander

1968 "Undergraduate achievement and institutional excellence." Science 161:661-668.

Becker, Gary

1964 Human Capital. New York: Columbia University Press.

Becker, Gary S.

1975 Human Capital. Second Edition. New York: Columbia University Press.

Becker, Gary and Barry R. Chiswick

1966 "Education and the distribution of earnings." American Economic Review 56:358-369.

Blau, Peter M. and Otis Dudley Duncan

1967 The American Occupational Structure. New York: Wiley.

Bowles, S.

1972 "Schooling and inequality from generation to generation." Journal of Political Economy 80:S219-S251.

Broom, Leonard, Paul Duncan Jones, F. Lancaster Jones, and Patrick McDonnell

1977 "Workers traits and worker functions in DOT." Journal of Vocational Behavior 11:253-261.

Burawoy, Michael

1977 "Social structure, homogenization and the process of status attainment in the United States and Great Britain." American Journal of Sociology 82:1031-1042.

Center for Human Resource Research

1976 The National Longitudinal Surveys. Columbus, Ohio: The Ohio State University.

Cohen, Malcolm S.

1971 "Sex differences in compensation." Journal of Human Resources 6:434-447.

Coleman, James

1966 Equality of Educational Opportunity. Washington, D. C.: United States Government Printing Office.

Coser, Lewis A.

1975 "Presidential address: Two methods in search of a substance." American Sociological Review 40:691-700.

Duncan, Otis Dudley

1961 "A socioeconomic index for all occupations." Pp. 109-138 in A. J. Reiss (ed.), Occupations and Social Status. New York: Free Press of Glencoe.

1969 "Inheritance of poverty or inheritance of race?" Pp. 85-110 in D. P. Moynihan (ed.), On Understanding Poverty. New York: Basic.

Featherman, David L. and Robert M. Hauser

1976 "Sexual inequalities and socioeconomic achievement in the U. S., 1962-1973." American Sociological Review 41:462-483.

1978 Opportunity and Change. New York: Academic.

Fuchs, Victor R.

1971 "Differences in hourly earnings between men and women." Monthly Labor Review 94:9-15.

Gourman, J.

1967 The Gourman Report. Phoenix, Arizona: The Continuing Education Institute.

Griffin, Larry

1976 "Specification biases in estimates of socioeconomic returns to schooling." Sociology of Education 49:121-139.

1978 "On estimating the economic view of schooling and experience: Some issues in conceptualization and measurement." Sociological Methods and Research 6:309-335.

Griffin, Larry and Karl Alexander

1978 "Schooling and socioeconomic attainment: High school and college influences." American Journal of Sociology 84: 319-347.

Griliches, Zvi

1976 "Wages of very young men." Journal of Political Economy 84: S69-S85.

Griliches, Zvi and William Mason

1972 "Education, income and ability." Journal of Political Economy 80:S74-S107.

Haley, William

1973 "Human capital: The choice between investment and income." American Economic Review 63:929-944.

Hines, Fred, Luther Tweeten and Martin Redfern

1970 "Social and private rates of return to investment in schooling, by race-sex groups and regions." Journal of Human Resources 5:318-340.

Jencks, Christopher

1979 Who Gets Ahead? New York: Basic.

Jencks, C., M. Smith, H. Ackland, M. Bane, D. Cohen, H. Gintis, B. Heyns, and S. Michelson

1972 Inequality: A Reassessment of the Effect of Family and Schooling in America. New York: Basic.

Kagan, Jerome

1964 "Acquisition and significance of sex typing and sex-role identity." Pp. 137-167 in M. Hoffman and L. Hoffman (eds.), Review of Child Development Research. New York: Russell Sage Foundation.

Karsh, Bernard

1976 "Industrial relations in Japan." Pp. 877-900 in R. Dubin (ed.), Handbook of Work, Organization, and Society. Chicago: Rand-McNally.

Kim, Jae-On

1975 "Multivariate analysis of ordinal variables." American Journal of Sociology 81:261-298.

King, Allan

1977 "Is occupational segregation the cause of the flatter experience-earnings profiles for women?" Journal of Human Resources 12:541-549.

Kish, Leslie

1965 Survey Sampling. New York: Wiley.

Kmenta, Jan

1971 Elements of Econometrics. New York: Macmillan.

Koch, James

1972 "Student choice of undergraduate major field of study and private internal rates of return." Industrial and Labor Relations Review 26:680-685.

Kohen, Andrew.

1973 "Determinants of early labor market success among young men." Ph.D. dissertation. Ohio State University.

Kohn, Melvin and Carmi Schooler

1973 "Occupational experience and psychological functioning: An assessment of reciprocal effects." American Sociological Review 38:97-118.

Labovitz, Sanford

1970 "The assignment of numbers to rank order categories." American Sociological Review 35:515-524.

Malkiel, Burton and Judith Malkiel

1973 "Male-female pay differentials in professional employment." American Economic Review 63:693-705.

Mincer, Jacob

1974 Schooling, Experience and Earnings. New York: Columbia University Press.

Mincer, Jacob and Solomon Polachek

1974 "Family investments in human capital: Earnings of women." Journal of Political Economy 82:S76-S108.

National Institute of Education

1977 Grants for Research on Education and Work. Washington, D. C.: U. S. Government Printing Office.

Nie, Norman, C. H. Hull, J. G. Jenkins, K. Steinbrenner, and D. H. Bent

1975 Statistical Package for the Social Sciences. Second Edition. New York: McGraw-Hill.

Oppenheimer, Valerie

1970 The Female Labor Force in the United States. Population Monograph Series, No. 5. Berkeley: University of California.

Polachek, Solomon

1975 "Potential biases in measuring male-female discrimination." Journal of Human Resources 10:205-229.

President of the United States

1977 Economic Report of the President. Washington, D. C.: U. S. Government Printing Office.

Renshaw, Edward

1960 "Estimating the returns to education." Review of Economics and Statistics 42:318-324.

Roby, Pamela

1975 "Structural and internalized barriers to women in higher education." Pp. 171-193 in J. Freeman (ed.), Women: A Feminist Perspective. Palo Alto: Mayfield.

Sawhill, Isabel

1973 "The economics of discrimination against women." Journal of Human Resources 8:383-395.

Shryock, Henry and Jacob Siegel

1973 The Methods and Materials of Demography. Volume 1. Washington, D. C.: U. S. Government Printing Office.

Stolzenberg, Ross

1975 "Education, occupation, and wage differences between white and black men." American Journal of Sociology 81:324-342.

1975 "Occupations, labor markets, and the process of wage attainment." American Sociological Review 40:645-665.

Suter, Larry and Herman Miller

1973 "Income differences between men and career women." American Journal of Sociology 78:962-974.

Temme, Lloyd

1975 Occupation: Meanings and Measures. Washington, D. C.: Bureau of Social Science Research.

U. S. Department of Labor

1965 Dictionary of Occupational Titles. Third Edition. Washington, D. C.: U. S. Government Printing Office.

1977 Dictionary of Occupational Titles. Fourth Edition. Washington, D. C.: U. S. Government Printing Office.

Wales, Terence J.

1973 "The effect of college quality on earnings: results from the NBER-Thorndike data." Journal of Human Resources 8:306-317.

Walther, Regis

1960 "The functional occupational classification project: A critical appraisal." Personnel and Guidance Journal 38:698-706.

Weitzman, Lenore

1975 "Sex-role socialization." Pp. 105-14 in J. Freeman (ed.) Women: A Feminist Perspective. Palo Alto: Mayfield.

Welch, Finis

1966 "Measurement of quality of schooling." American Economic Review 56:379-392.

Women's Bureau

1965 Handbook on Women Workers. Bulletin 294. Washington, D. C.: U. S. Government Printing Office.