DOCUMENT RESUME

ED 063 686 EA 004 427

AUTHOR Golladay, Fredrick L.; And Others
TITLE Education and Distribution of Income: Some Exploratory Forays. Technical Reports. Conference on Policies for Educational Growth (Paris, France, June 3-5, 1970).
INSTITUTION Organisation for Economic Cooperation and Development, Paris (France).
PUB DATE 20 May 70
NOTE 164p.; Related to Background Study No. 11

EDRS PRICE MF-$0.65 HC-$6.58
DESCRIPTORS Cost Effectiveness; *Education; *Educational Benefits; *Educational Economics; Educational Finance; *Educational Policy; Educational Research; Educational Sociology; Equal Education; Higher Education; *Income; Investment

ABSTRACT

The first paper in this series, "Problems in the Econometric Analysis of Educational Technology," by Fredrick L. Golladay, examines the analogy between empirical production functions in economic research and the school characteristic study. The second paper, "Problems in Making Policy Inferences from the Coleman Report," by Glen Cain and Harold Watts, presents the theoretical bases for causally interpretable multivariate empirical research into education. The third paper, "Education and Income: A Study of Cross-Sections and Cohorts," by Robinson Hollister, summarizes research into the problems of conceptualization and the measurement of school characteristics, inputs, and outputs. The fourth paper, "Patterns of Rates of Return to Investment in Education: Some International Comparisons," by W. Lee Hansen, reviews briefly the statistical problems that emerge in school characteristic studies. The final paper, "The Search for Equity in the Provision and Finance of Higher Education," by W. Lee Hansen and Burton A. Weisbrod, summarizes the discussions and draws conclusions for the appropriate focus of research into resource allocation to and within education. Related documents are ED 057 470 and EA 004 323. (Author)

CONFERENCE ON POLICIES FOR EDUCATIONAL GROWTH

PARIS, 3-5 JUNE 1970

EDUCATION AND DISTRIBUTION OF INCOME

VII

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT PARIS 1971

FOREWORD

The 1970 Paris Conference on Policies for Educational Growth was organized by OECD as a sequel to its 1961 Washington Conference on Economic Growth and Investment in Education. The purpose of the Conference was to assess the nature and consequences of the expansion of education in OECD countries during the last 10-15 years and to foresee the main policy problems arising from continued educational growth in the future.

The proceedings of the Conference are presented in a set of eight volumes consisting of:

- The General Report of the Conference published under the title: EDUCATIONAL POLICIES FOR THE 1970's,

and the following series of documents containing the twelve supporting studies prepared by the Secretariat:

II - EDUCATIONAL EXPANSION IN OECD COUNTRIES SINCE 1950 - (Background Report No. 1).

III - TRENDS IN EDUCATIONAL EXPENDITURE IN OECD COUNTRIES SINCE 1950 - (Background Report No. 2).

IV - GROUP DISPARITIES IN EDUCATIONAL PARTICIPATION AND ACHIEVEMENT:

Group Disparities in Educational Participation - (Background Report No. 4).

Differences in School Achievement and Occupational Opportunities - Explanatory Factors. A Survey based on European Experience - (Background Report No. 10).

V - TEACHING RESOURCES AND STRUCTURAL CHANGE:

Teaching Staff and the Expansion of Education in Member Countries since 1950 - (Background Report No. 3).

Changes in Secondary and Higher Education - (Background Report No. 6).

Educational Technology: Practical Issues and Implications - (Background Report No. 7).

VI - THE DEVELOPMENT OF EDUCATIONAL PLANNING:

Educational Policies, Plans and Forecasts during the Nineteen-Sixties and Seventies - (Background Report No. 5).

Educational Planning Methods - (Background Report No. 8).

The Role of Analysis in Educational Planning - (Background Report No. 9).

VII - EDUCATION AND DISTRIBUTION OF INCOME - (Background Report No. 11).

VIII - ALTERNATIVE EDUCATIONAL FUTURES IN THE UNITED STATES AND IN EUROPE: METHODS, ISSUES AND POLICY RELEVANCE - (Background Report No. 12).

CONTENTS

II

TECHNICAL REPORTS RELATED TO BACKGROUND REPORT No 11

CONTENTS

Paper 3

EDUCATION AND INCOME: A STUDY OF CROSS-SECTIONS AND COHORTS

Paper 4

PATTERNS OF RATES OF RETURN TO INVESTMENT IN EDUCATION: SOME INTERNATIONAL COMPARISONS

Paper 5

THE SEARCH FOR EQUITY IN THE PROVISION AND FINANCE OF HIGHER EDUCATION

Paper 1

PROBLEMS IN THE ECONOMETRIC ANALYSIS OF EDUCATIONAL TECHNOLOGY

by

Fredrick L. Golladay[1]

1. The author wishes to acknowledge the financial support of the Institute for Research on Poverty, University of Wisconsin, and the Organisation for Economic Co-operation and Development. He also wishes to thank W. Lee Hansen, Robinson Hollister, Louis Emmerij and M. Vere DeVault for helpful comments on earlier drafts of the paper. Marilyn Hastings and Mary Alberston Golladay have provided considerable research assistance.

INTRODUCTION

The importance of education to various dimensions of national development and to individual welfare has been thoroughly documented. However, to provide educational opportunities as abundantly as educationists and social scientists would urge is beyond the economic resources of even the richest nation. Increasing the efficiency with which educational resources are allocated is therefore imperative.

A number of recent studies have attempted to provide insights into the educational process and into the efficiency of resource allocation to and within education through multivariate statistical analyses of data drawn from schools and school systems. The ultimate objective of these studies has been to evaluate the strength of causal relationships between school inputs and characteristics and student achievement. Several scholars have challenged the findings of these studies; despite rather widespread criticism of school characteristics studies, it is apparent that the conclusions of these projects are influencing educational policy. Not only have the results of these studies been cited as justification for policy regarding resource allocation to and within educational systems, but attempts have also been made to discount criticisms of the studies. One well-known policy adviser has even suggested that the objections raised by critics are not legitimate but rather are attempts to preserve some cherished and long-held preconceptions regarding the educational process[1].

The purpose of this paper is to provide a comprehensive, systematic examination of the theoretical, conceptual and empirical problems plaguing research into the educational process which is based upon extensive observational data. The goal of this discussion is to demonstrate that the shortcomings of school characteristics studies are so pervasive and fundamental as to undermine any confidence that one might place in the findings of such research. An extensive critical literature has emerged which in somewhat fragmentary fashion has attacked aspects of particular studies; most of the discussion has focused upon empirical defects of the studies. This paper addresses a more basic issue - whether the large-scale, observational study of education is an appropriate research strategy for examining the technology of the educational process.

This paper is presented in five parts. Part I examines the analogy between empirical production functions in economic research and the school characteristics study. Several scholars have explicitly drawn the analogy and have borrowed heavily upon the more developed literature of production function estimation in preparing research strategies. This discussion considers the implications of public production of educational services for the estimation and interpretation of production functions for education. Part II presents the theoretical bases for causally interpretable, multivariate empirical research into education. Observational studies require a well-developed, theoretical system in order to identify relevant variables and to suggest appropriate functional relationships among dependent and independent variables. This part of the paper both reviews the need for theory in empirical research of the type being considered here and assesses the literature of educational theory in an attempt to construct an empirically relevant, operational-theoretical model of the educational process. Part III summarizes

1. Daniel P. Moynihan, "Sources of Resistance to the Coleman Report", Harvard Educational Review, 38, 1, 1968, pp. 23-36.

research into the problems of conceptualization and measurement of school characteristics, inputs and outputs. Attention is focused upon the absence of well-scaled, operational measures of these variables and the resulting implications for analysis and interpretation of empirical relationships. Part IV briefly reviews the statistical problems that emerge in school characteristics studies. Part V summarizes the discussions and draws conclusions from the paper for the appropriate focus of research into resource allocation to and within education.

It should be noted that each part of the paper has been presented as though the problems considered in the other parts did not exist. This has been done in order to simplify the discussions; the reader should bear in mind that the discussion of a problem which occurs whether or not other defects have been corrected does not imply that the author has dismissed the earlier defects.

I

PRODUCTION FUNCTION ESTIMATION FOR PUBLIC SECTOR ACTIVITIES

The production function concept has been demonstrated to be a useful empirical tool with which to derive insights into efficient responses of private firms to changes in factor or product prices. Several scholars have suggested that application of production function concepts to the analysis of technical alternatives in public education would be fruitful[1]. However, several critical assumptions employed in production function estimation are violated by public production activities. The purpose of this part is to indicate the nature of these assumptions and to underscore the consequences of their being violated in studies of educational practices.

The usefulness of production functions in the study of the behaviour of private firms depends crucially on the assumption of maximizing behaviour. The assumption of profit maximization permits one to assert that observed production techniques are efficient; the theory of the firm offers compelling evidence that enterprises which are not efficient in the economic sense will not survive the rigours of competition. Economic efficiency, defined as minimization of the cost of production, in addition implies that techniques being employed are efficient in a more narrow physical or engineering sense. The variations observed in technology may therefore be assumed to reflect optimal adjustments to changes in economic conditions, particularly regarding the cost of factor services and materials.

The assumption of profit maximization also implies that factors will be paid approximately their marginal contributions to output; this implication is important because it facilitates the aggregation of heterogeneous inputs into production. The broad category of inputs denoted as labour may be usefully analysed by a single dimensional index expressed as dollars worth of labour productivity, if one may assume that marginal productivity theory is appropriate. In the absence of this highly useful implication of maximizing behaviour, one is forced to regard every distinguishable skill as an input to be analysed explicitly.

It is perhaps obvious that the implications of technical efficiency and consistent aggregation of inputs derived from the assumption of maximizing behaviour are not appropriate to studies of production in the public sector generally. This conclusion is particularly important to research into educational technology based upon observational data.

Educational decision makers do not appear to maximize any well-defined criteria function. Purists may argue that educational decisions, if rational and consistent, must reflect some underlying set of objectives which are being optimized, at least implicitly. One may concede this point, yet recognize that the decentralized structure of educational systems creates a presumption that the latent decision functions of teachers, school administrators and public officials are likely to be highly diverse. The

1. For example, see Samuel S. Bowles, "Towards An Educational Production Function", in National Bureau of Economic Research Conference on Income and Wealth, New York, 1968.

observed technical choices may be optimal with respect to an almost infinite variety of objectives, and there is no reason to expect these objectives to be mutually consistent or reinforcing.

The absence of a well-defined objective function which is being maximized by educational decision makers destroys the implication of production theory, that observed technologies are efficient in an engineering sense. One cannot argue that observed techniques are efficient methods for obtaining a particular output where even the output being maximized is ambiguous.

The second implication of profit maximizing behaviour, that marginal productivity theory explains payments to factors, is also inappropriate to education. The highly convenient procedure of aggregating inputs by values assigned by the market is therefore not legitimate. It may be argued that the opportunity cost of remaining in the teaching profession is determined by market forces; profit maximizing firms have determined the value of a teacher as a production worker. This suggests, however, only that a lower bound on the value of a teacher is established; even this value would appear to have little relevance to one's productivity as a teacher. The widespread use of salary schedules, seniority increments and the reluctance to employ "merit" salary payments reinforces the view that marginal productivity theory is not an adequate conceptualization of wage determination in education.

Similar arguments demonstrating the inappropriateness of aggregating inputs by using market valuation may be made for classes of educational inputs other than teachers. There is no a priori basis for the view that the value of educational resources bears any consistent relationship to a well-defined, general index of school outputs. For example, lavish athletic facilities may contribute to a community's sense of prestige without having any positive impact on student reading ability. Educational inputs must therefore be considered in a highly detailed and disaggregative fashion if one is to capture the variability in school inputs.

Implicit in much of the above criticism of production function estimation is the realization that the outputs of formal education are as elusive as school inputs. The issue of school decision-making might be reinterpreted as an examination of the question, "What do schools produce?". Because education is typically provided publicly without direct charge, it is impossible to infer a well-defined, general index of school outputs. The outputs of many industries are equally diverse; however the marketing of these goods provides a rigorous evaluation of the social values attached to these goods, permitting one to estimate a single dimensional measure of production. In short, market prices may be used as aggregation weights for related but heterogeneous outputs. A community's preferences for advanced placement physics education, remedial reading instruction and inter-school athletics are never subjected to these rigours of market evaluation.

This part of the paper has attempted to demonstrate that empirical production functions for education estimated from observational data confront important problems because of the nature of public sector activities. The interpretation of the estimated production function for education as a summarization of technically efficient alternatives is destroyed by the absence of consistent maximizing behaviour; the estimated function summarizes a variety of often unrelated production activities. In addition, there is no theoretical basis for the assumption that the estimated relationships are in any sense technically efficient. Finally, the absence of an elegant theoretical justification for aggregation of either school inputs or outputs considerably complicates the study of educational technology. Part IV considers the aggregation problem and related statistical issues in greater detail.

II

EDUCATIONAL THEORY AND EMPIRICAL PRODUCTION FUNCTION ESTIMATION FOR EDUCATION

The purposes of this Part are to indicate the requirements for causally interpretable, non-experimental research into educational technology, and to provide some initial suggestions regarding the appropriate strategy for such research. First, the relationship of theory to empirical research is considered; it is stressed that non-experimental or observational research requires a theoretical framework with which to identify important variables and to specify the algebraic form of the relationships. Second, the theoretical literature of education is examined in an attempt to develop a tentative model of the educational production process. The discussion is organised into investigations of the inter-temporal structure of education, the theory of classroom learning and the theory of educational administration. The recursive structure of student development is supported both with speculative analyses and empirical studies. The consideration of classroom learning departs from several earlier studies in that it is not explicitly concerned with abstract learning theory but rather focuses upon instruction. The discussion of administrative theory is designed to illuminate educational decision-making, particularly with respect to the extent to which educational developments are innovated.

The Role of Theory in Empirical Research

The dangers of adopting a strictly empiricist research strategy have been widely discussed in the literature of sociometric and econometric research[1]. The discovery of a systematic relationship between two variables suggests several hypotheses. The two variables may be causally related, although it is uncertain in which direction the causation might operate, or, indeed, if a unique direction of causation may be identified; examples of mutual causation abound in the study of social systems. Alternatively, the two variables may share a common source of causation or may simply measure the same latent concept. In the absence of a theory of the process to guide in the specification of functions to be estimated, the results of statistical analyses may only be regarded as predictive as opposed to causally interpretable. The empiricist study describes phenomena in terms of measurable variables and may enable one to predict a more elusive variable through the use of readily available measures; such a study does not, however, provide an empirical basis for conclusions regarding the consequences of manipulating particular variables. The policy oriented study of educational technology being considered here requires that the results of the research be causally interpretable and that one be able to evaluate the consequences of manipulations of school input variables.

The statistical technique of multiple regression analysis produces estimates of the parameters of a linear function such that the estimated values of the dependent variable are as highly correlated as possible with observed values of the dependent variable. The technique by itself does not generate or

1. One of the best statements regarding the potential dangers of the strictly empiricist approach is found in Tjalling J. Koopmans, "Measurement Without Theory", Review of Economics and Statistics, August 1947, pp. 161-172.

test causal hypotheses, nor does it provide a rigorous basis for selecting variables for analysis or functional relationships. A number of scholars have demonstrated that the selection of variables and specification of appropriate functional forms must be based upon a priori and deductive theoretical analyses of the process under study. Regression analysis may then be appropriately used to estimate the strength or magnitude of the association of variables.

Successful generalization of the production function concept to the study of educational technology thus requires that an adequate theory of the educational process be obtained. Most empirical research into educational technology has simply identified the variables which casual empiricism would suggest are relevant and proceeded to assume that the computationally-expedient, linear form is appropriate. The idea that educational achievement is simply the sum of the independent influences of a number of school inputs is intuitively absurd. Recently, it has been suggested that the field of learning theory might provide insights into the appropriate specification of educational production functions.

Educational Theory and the Modelling of the Educational Process

This section offers an approach to the specification of educational theory that would furnish the necessary basis for production function estimation. A review of theoretical discussion and empirical research relating to educational processes suggests variables which are appropriate to a quantifiable model of the educational process. A detailed model of the educational process would be both beyond the scope of the present paper and premature, given existing knowledge; the section is thus meant to be sugges''ve only.

This discussion differs markedly from earlier efforts in that it considers educational processes from the viewpoint of the literature of educational theory and research. Previous efforts at constructing quantifiable models of education have been characterized by their conspicuous lack of a basis drawn from educational literature. Intuitive justification for a model composed on the basis of empirical expediency has been offered by reference to a small number of empirical studies and theoretical summaries. While such efforts are notable in their intent, approaching the problem of estimating quantitative relationships in education without an understanding of educational processes increases the danger of misspecification of underlying principles and consequent misinterpretation of empirical results.

This discussion will be divided into two subsections. First, the general structure of the educational process will be studied to illuminate the issue of an appropriate choice of model structure. Second, the theory of education and of the administration of education will be perused to obtain insights into the classes of variables which affect the educational process. Following this section, the measurement and conceptualization of variables will be discussed to suggest approaches to securing data and facilitating the estimation of model parameters.

a) Structure of the Educational Process

Education is a sequential process in which a learner's prior knowledge is affected (increased, altered, or eliminated) by new experiences. The outputs from one stage of the process serve as inputs to another. Thus, education may be viewed in a recursive framework which represents a dynamic process as a series of sequential stages, each one dependent upon past stages[1]. This approach incidentally coincides with the design of formal educational structures which employ grades, levels or types of schooling.

The view of education as a recursive system is consistent with existing knowledge of intellectual development, the learning process and the institutional structure within which education takes place. At the most naive level, it is apparent that mastery of prerequisite, basic skills is crucial to future

1. H.O.A. Wold, "A Generalization of Causal Chain Models", Econometrica, 28, 2, April 1960, pp. 443-463.

academic success: basic verbal, reading and arithmetic skills are necessary inputs into more advanced literary or quantitative subjects. Less obvious is the exact nature of the role of prerequisite knowledge in intellectual development. Work by the pioneering psychologist, Jean Piaget, indicates that individuals must pass through three stages of intellectual development with respect to any subject[1]. First, one accumulates concrete knowledge about the subject; second, one develops generalizations based upon concrete experiences or knowledge; finally, one resorts to abstract analysis of the important aspects of events or phenomena. Abstractions are meaningless in the absence of general and concrete referents.

The importance of prerequisite knowledge to student learning has been amply supported by empirical research. Research studies indicate that a student is not able to perform a "higher level" learning task if he has not been able to perform a "lower level", or prerequisite task[2]. Where prerequisite tasks were taught and mastered at some level, the success rate at learning new tasks far exceeded that rate which is normally observed in educational practice. The widespread research procedure of statistically controlling for previously measured abilities in studying effects of curriculum materials or instructional techniques demonstrates implicit recognition of the role of prior knowledge as an input in the educational process.

Grade levels, cycles or forms characterize the institutional structure of educational systems in many countries. While orginally an expedient resulting from the need to meet the demands of widespread education with limited resources[3], the structure has continued because of its convenience as well as its economic rationality. Many recent attempts at individualized instruction[4], and non-graded schools[5], reveal a trend which may eventually replace the rigid institutional patterns of the past. While this trend will have long-range implications for research into educational processes, some of the practical, logistical problems posed by attempts at genuinely individualized instruction indicate that in the near future educational researchers will for the most part be working with data generated from graded schools.

A recursive model of education has the practical advantage of permitting one to examine the efficiency of school resources committed to various levels and subject areas of the educational programme. Empirical analysis might, for example, indicate that development of reading skills should be given greater emphasis at the expense of arithmetic skills in early grades. Furthermore, a recursive model permits one to examine the performance consequences of inferior educational experiences at each grade level. The appropriate level of disaggregation over time and subject matter is considered below in the discussion of conceptualization and measurement of variables, and again in the review of statistical estimation procedures.

b) Theory of the Educational Process

Educational theory provides insights into the identification and delineation of factors affecting educational achievement. An attempt to determine the appropriate variables and the nature of their influence suggests that two types of theories must be examined - educational theory and administrative theory. Educational theory describes the nature of the educational process and the manner in which learning occurs. Administrative theory indicates the way in which decision-making affecting the allocation of

1. John H. Flavell, The Developmental Psychology of Jean Piaget, Princeton, New Jersey: D. Van Nostrand Company, Inc., 1963.

2. Robert Gagne, "The Acquisition of Knowledge", Psychological Review, 69, 1962, pp. 355-365; and Gagne, J.R. Mayor, H.L. Garstens and N.E. Paradise, "Factors in Acquiring Knowledge of a Mathematical Task", Psychological Monographs, 76, 1962, Vol. 7, Whole No. 526.

3. R.L. Butts and L.A. Cremin, A History of Education in American Culture, New York: Henry Holt, 1959.

4. "Association for Supervision and Curriculum Development," Individualizing Instruction, Washington: the Association, 1964, and "National Society for the Study of Education", Individualizing Instruction, Chicago: University of Chicago Press, 1962.

5. John Goodlad and Robert H. Anderson, The Nongraded Elementary School, New York: Harcourt, Brace and World, 1963

resources to and within education is conducted. Thus, one provides the technical basis for the production function and one describes the "market" within which education occurs. The following discussion will treat first educational theory and then administrative theory.

i) Educational theory

Variables influencing the educational process may be considered within three clases. Exogenous variables are defined as those external to the educational system and hence beyond the direct control of educational decision makers. Endogenous variables are those input variables which may be manipulated by the educator and which are aspects of the educational process. The third class of variables includes educational outputs. This conceptualization of education variables is somewhat artificial in view of the dynamic nature of education; exogenous inputs into schools are augmented and transformed into school outputs which are later useful as endogenous inputs. Nonetheless, this classification of variables clarifies our discussion of educational theory.

The relationship among these three classes of variables may be summarized as:

$$E_t = F(E_{t-1}, X_t, N_t)$$

where E_t = output for some measurable aspect of school effect for period t;

X_t = input factors exogenous to the formal educational system;

N_t = input factors endogenous to the educational system.

Each of these three classes is, in turn, made up of several sets of variables, which may be identified as:

Variables within E:
- C = cognitive skills,
- R = affective responses,
- D = physical development.

Variables within X:
- A = aptitude: genetic endowment supplemented by experience,
- P = prerequisite knowledge: verbal, analytic and substantive skills necessary for academic success,
- M = motivation: activity directed or sustained towards a scholarly goal.

Variables within N:
- S = schooling process: curriculum, instructional method, teacher expertise,
- L = classroom: teacher interaction with students, peer culture,
- F = school environment: facilities.

The following discussion will suggest a theoretical basis for the inclusion of each set of variables in a model of the educational process and will provide a summary discussion of related research findings. Space does not permit an exhaustive review of educational research, nor the synthesis of the factors identified here into an elaborate model. The purpose of this discussion is to illuminate, for persons outside of the professional sphere of the educationist, critical factors which should be considered by all scholars attempting research into the effects of education.

The present discussion differs from previous approaches to the identification of variables significant in studying education in at least one major aspect: socio-economic and cultural variables are not included in the discussion per se. In excluding them, the author is not ignoring the significant correlative

findings linking educational achievement with class or socio-economic status. Many factors influencing educational outcomes which may reflect cultural characteristics are noted in the following discussion. However, attempts have been made to keep the variables identified here as abstract as possible. The use of socio-economic characteristics as proxy variables for more basic traits important to education is highly subject to criticism and will be explored in detail below.

The important issue of obtaining suitable measures of variables arises for almost every variable treated here. Therefore, related questions and their implications will be examined in the next part of the paper. Specific measurement problems will be mentioned only if they illuminate an important aspect of the use of the variable in empirical research.

Educational output (E)

An extensive literature on the goals of education suggests the heterogeneity of output which is frequently desired as the result of schooling[1]. Of the three domains[2] of educational output which are presented here (cognitive skills (C), affective responses (R), physical development (D)), the development of the first two is most likely to be dominated by the schools and, hence, they will constitute the formal output of schools for purposes of the present discussion.

The two areas of cognitive skills and affective responses have been treated separately in taxonomies which list educational objectives according to hierarchies of skill and abstraction required for their mastery. Yet, in the recursive educational process, the achievement of an objective in one domain may act as a prerequisite or motivating factor for subsequent achievement and, hence, distinctions between the two domains are not always evident. No attempt has been made to separate the domains into disjoint categories in the following discussion of educational inputs. Specific input factors are examined as they contribute to output identified with one or the other or both domains.

Educational inputs: exogenous factors (X)

While student characteristics, the exogenous factors, may be regarded as the raw inputs into the educational production process, because of the recursive structure of schooling, these variables are purely exogenous factors only in the initial period. After having been operated upon by the educational process, the analytic distinction between exogenous and endogenous becomes somewhat blurred. In the following discussion, exogenous factors are presented in such a way as to suggest the impact that schooling may have upon them.

Aptitude (A)

It is easy to verify that individuals of the same age rarely at any given time are able to perform exactly the same sets of tasks. This characteristic, commonly called intelligence or aptitude (the two terms are used interchangeably here), affects scholastic performance. This attribute is not constant from birth[3] but, rather, represents a developmental characteristic[4]. The intelligence quotient

1. For example, see Rockefeller Brothers Fund, "The Pursuit of Excellence: Education and the Future of America", Panel Report V of the Special Studies Project, Garden City, New York: Doubleday, Inc., 1958; J.W. Gardner, "National Goals in Education", Goals for Americans, the Report of the President's Commission on National Goals, Englewood Cliffs, New Jersey: Prentice-Hall, 1960; and C.M. Lindvall (Ed.), Defining Educational Objectives, Pittsburgh: University of Pittsburgh Press, 1964.

2. From B.S. Bloom, A Taxonomy of Educational Objectives: Cognitive Domain, New York: David McKay, Inc., 1965; and David Krathwohl, A Taxonomy of Educational Objectives: Affective Domain, New York: David McKay, Inc., 1964.

3. J.P. Guilford presents evidence refuting this widely-held principle in The Nature of Human Intelligence, New York: McGraw-Hill Book Company, 1967. See also J.M. Hunt, Intelligence and Experience, New York: Ronald Press, 1961.

4. B.S. Bloom, Stability and Change in Human Characteristics, New York: John Wiley and Sons, Inc., 1964.

(IQ) compares actual achievement of an individual at a given age with that of a normed population. The measure includes the influence of environment on a person's knowledge at any point in time. Thus, when used as an empirical variable in research, intelligence is an intermediate concept incorporating principles of genetic endowment and environmental influence in the form of past experience.

In addition to showing that intelligence as measured by IQ is a variable factor[1], research also suggests that the influence of environment upon intelligence is greatest in the early years of a child's development. Particularly, lack of learning in one period is difficult, sometimes impossible, to make up in later periods[2]. Measured differences in intelligence which previously were attributed to ethnic or class intelligence differences may reflect this modifying influence of environment[3].

Prerequisite knowledge (P)

Prerequisite knowledge may be considered as a specific extension of the environmental component of aptitude. It encompasses academic abilities exclusive of social and environmental factors. The importance of prerequisite knowledge to student achievement in learning tasks has already been mentioned to lend support to the consideration of education as a recursive process[4].

The role of prerequisite learning in various disciplines can be expected to vary significantly, depending on the technical nature of the material being studied and the need for highly specific principles or tasks as prerequisites. This, in turn, would suggest that an educational production function should be disaggregated by subject, as the role of prerequisites in one subject (e.g. mathematics) may be considerably different from that in another (e.g. fine arts or literature). In instances where early prerequisite knowledge was missing, and where it was highly important, the academic advancement of a student would be expected to decrease rapidly and approach zero, unless the academic situation were one which would allow instruction at the level of the student's need. In short, the problems of poor achievement may be expected to compound, suggesting that the relationship between adequate prerequisites and academic success is non-linear.

Motivation (M)

The definition of motivation as "the combination of forces which initiate, direct or sustain activity towards a scholarly goal[5] ", implies that motivation is a composite of qualities including expectation and reinforcement. While theoretical discussions of motivation are ambiguous, or even contradictory, it is apparent that some minimal amount of motivation must be present before learning can take place[6]. Beyond that, theoretical works are in conflict. It has been asserted that performance increases monotonically as motivation increases[7], implying that the most effective instruction occurs when motivation is maximal. A more plausible argument is that before motivation can facilitate performance, correct

1. Extensive evidence to support this conclusion is provided in Robert L. Green and others, The Educational Status of Children in a District without Public Schools, Washington, D.C.: Department of Health, Education and Welfare, United States Office of Education, Cooperative Research Project No. 2321, 1964; Otto Klienberg, Negro Intelligence and Selective Migration, New York: Columbia University Press, 1935; E.S. Lee, "Negro Intelligence and Selective Migration: a Philadelphia Test of the Klienberg Hypothesis", American Sociological Review, XVI, 1951, pp. 227-233.
2. B.S. Bloom, Stability and Change in Human Characteristics, New York: John Wiley and Sons, Inc., 1964.
3. Martin Deutsch and Bert Brown, "Social Influences in Negro-White Intelligence Differences", Journal of Social Issues, April 1964, pp. 24-35.
4. Robert Gagne, "The Acquisition of Knowledge", op. cit.
5. William W. Farquhar, "Academic Motivation and Inner-City Schools", in Herbert C. Rudman (Ed.), Urban Schooling, New York: Harcourt, Brace and World, Inc., 1968, p. 198.
6. John W. Atkinson, An Introduction to Motivation, Princeton, New Jersey: Van Nostrand, 1964.
7. C.E. Osgood, Method and Theory in Experimental Psychology, New York: Oxford University Press, 1953, p. 413.

or otherwise desired behaviour must be dominant over other possible behaviour patterns[1]. In other words, motivation can be disfunctional with respect to academic achievement if habits which facilitate learning do not exist. Research results on the effects of motivation upon academic achievement also suggest that achievement motivation in a given situation is dependent on both the individual's typical level of motivation and the degree to which he sees the situation as achievement relevant[2]. Thus, a student must be able to see the relationship of a particular task to an achievement goal.

Educational inputs: endogenous factors (N)

Endogenous factors are those over which educational policy makers have the most immediate control, namely, school characteristics. Studies of the influence of these characteristics on educational production have special significance for the issues of allocation of resources to and within the educational system.

Schooling process (S)

Curriculum quality, instructional method and teacher skill are important in educational production. Schooling process variables which deal with these aspects of educational practice measure the efficiency of the educational system. Educational research is abundant in these three areas. No doubt part of the interest in research of this nature has been stimulated by the widespread educational reform which has occurred in the past decade in many subject areas in a great number of countries. The three classes of schooling process variables cited here - curriculum organisation, instructional method and teacher skill - will be discussed separately.

The majority of research studies in curriculum have utilized formal research designs to study a particular problem at a micro level. Results from these studies suggest that student achievement is related to curriculum objectives. As an example, many experiments have tested the efficiency of "new" versus "old" mathematics curriculum materials. In instances where subjects are tested with instruments designed on the basis of each set of materials, subjects using traditional materials typically perform best on tests designed for traditional programmes stressing mechanical skill in problem solving, and subjects using new materials perform best on new tests designed to measure analytic abilities[3]. These results imply that a critical aspect of the study of educational production functions is the choice of an objective function made with reference to the cultural, political, social and economic needs of the students in a particular educational system.

Research in instructional methods has illuminated psychological principles of learning efficiency and the influence of ability factors on learning patterns. It is not the purpose of the present summary to present in detail the research findings which have contributed to an understanding of student learning. However, some results have implications for the structuring of educational production research studies and will be summarized here.

Numerous studies suggest interactions of ability and the effect of knowledge of results upon achievement. If students' knowledge of the success of their academic performance is immediate, there appears to be no correlation between ability as measured by IQ and performance, or between reading level and performance[4]. When knowledge of results is less immediate, both IQ and reading level correlate with

1. K.W. Spence, Behavior Theory and Conditioning, New Haven: Yale University Press, 1956.
2. Elizabeth G. French, "The Development of a Measure of Complex Motivation", United States Personnel Training Research Center Research Report, No. 56-58, 1956; and French and I. Chadwick, "Some Characteristics of Affiliation Motivation", Journal of Abnormal Social Psychology, No. 52, 1956, pp. 296-300.
3. Milton Maier, "Evaluation of a New Mathematics Curriculum", American Psychologist, No. 17, 1962, p. 336.
4. L.D. Eigen, "A Comparison of Three Modes of Presenting a Programmed Instruction Sequence", Journal of Educational Research, No. 55, 1962, pp. 453-460; and J.K. Little, "Results of Use of Machines for Testing and for Drill Upon Learning in Educational Psychology", Journal of Experimental Education, No. 3, 1934, pp. 45-49.

performance on immediate and delayed post-tests. In addition to suggesting that the use of reinforcement through immediate feedback can be expected to have greater effect on poor students than on superior ones, these results suggest that the impact of instructional method on students of different abilities should not be uniform.

The growing use of a variety of media for instructional purposes has raised many questions regarding both the pedagogical effectiveness and the economic efficiency of mediated instruction. Abundant research in the field of instructional media has been carefully summarized[1]. Research indicates essentially that teachers and media can both be used for effective instruction. When extraneous factors are removed from an experiment in instructional effectiveness, significant differences almost never appear[2]. At the same time, when a comparison of methods focuses on a task for which a teacher is well-suited, classroom technique appears more effective. When a comparison considers a subject which requires particularly detailed, mediated presentation, the media is likely to appear most effective[3]. These results indicate that the educational policy maker has available numerous means of meeting particular goals. They also suggest that economic criteria may well be highly relevant at a time when resources for education are at a premium. Unfortunately, almost no research has examined the issue of expenditure allocation within schools to achieve optimal results from staffing patterns, technology and other facilities[4].

Teacher skill influences educational outcomes. Teachers have considerable influence over what transpires in the learning process as they manipulate aspects of the instructional situation: a teacher may control the stimulus situation, the verbal communication used to direct the learner and the positive or negative feedback to the students from the events of instruction[5]. The feedback to students acts as positive and negative reinforcement and affects academic performance directly by guiding students' progress and indirectly by affecting motivation levels[6].

Theoretical studies and empirical research on teaching methods to facilitate learning have related teacher effectiveness to learning principles. However, in spite of theoretical constructs outlining the nature of teacher effectiveness, research attempting to measure teacher effectiveness has produced little concrete evidence regarding the factors which are responsible for teacher quality[7]. Furthermore, no fully satisfactory measures for teacher quality have been found. Commonly used variables such as teacher verbal aptitude, experience and academic background do not provide consistent results[8]. These variables probably act as proxies for more basic and relevant variables which remain elusive.

1. For comprehensive bibliographies see, W. Schramm, "Learning from Instructional Television", Journal of Educational Research, No. 32-2, April 1962, pp. 156-157; Robert Glaser (Ed.), Teaching Machines and Programmed Learning, Washington, D.C.; National Education Association, 1965; M.A. May and A.A Lunsdaine, "Mass Communication and Educational Media", Annual Review of Psychology, 1965, pp. 475-534; and B.C. Duke (Ed.), Survey of Educational Media Research in the Far East, Washington, D.C., United States Office of Education, 1963.

2. D.W. Stickels, A Critical Review of the Methodology and Results of Research Comparing Television and Face-to-Face Instruction, unpublished doctoral dissertation, Pennsylvania State University, 1963.

3. Wilbur Schramm, Philip H. Coombs, et al., The New Media: Memo to Educational Planners, Unesco International Institute for Educational Planning, 1967.

4. Bruce R. Joyce, "Staff Utilization", Review of Educational Research, Vol. XXXVII, No. 3, June 1967, pp. 323-336.

5. Robert M. Gagne, The Conditions of Learning, New York: Holt, Rinehart and Winston, Inc., 1965.

6. E.B. Page, "Teacher Comments and Student Performance", Journal of Educational Psychology, Vol. XLIX, 1958, pp. 173-181; and Pauline S. Sears and Ernest R. Hilgard, "The Teacher's Role in the Motivation of the Learner", in Theories of Learning and Instruction, 63rd Yearbook, National Society for the Study of Education, Chicago: the Society, 1964, pp. 182-209.

7. For example, see A.S. Barr, et al., "Wisconsin Studies of the Measurement and Prediction of Teacher Effectiveness: A Summary of Investigations", Journal of Experimental Education, 1961, pp. 1-155.

8. One of the more extensive studies is that by David G. Ryans, Characteristics of Teachers: Their Description, Comparison and Appraisal, Washington, D.C.: American Council on Education, 1960. Other studies, including that by J.A. Shea, reveal very low correlations between measures of teacher effectiveness and mental ability. ("The Predictive Value of Various Combinations of Standardized Tests and Sub-tests for Prognosis of Teaching Efficiency", Catholic University of America, Educational Research, Monographs, 19, 1955, No. 6).

Classroom factors (L)

Learning is not exclusively influenced by cognitive considerations. Teacher personality and peer group attitudes affect motivation, and hence achievement, while a student is in an academic situation. These significant classroom factors operate simultaneously with the learning process.

Several aspects of teacher personality have an impact on classroom climate. The extent to which the teacher offers an image which the students regard as an "ideal" may affect their willingness to perform academic tasks. This may in part explain the difficulties typically encountered by middle-class teachers reared in suburban areas when they attempt to teach ghetto children[1]. In addition to the effects of different cultural environments of students and teachers on academic achievement, the interaction of teacher personality and student personality factors frequently acts to encourage or inhibit academic progress.

Research on the effect of teacher behaviour over the cycle of learning units and sub-units has provided insights into the effect of teacher behaviour upon student achievement.

There is a direct relationship between a teacher influence which encourages student participation and constructive pupil attitudes towards the teacher[2]. Research specifically studying the relationship between teacher influence and student achievement suggests that attitudes and achievement of students are superior when classroom teachers are able voluntarily to vary their influence upon the class during the learning cycle and are able to diagnose student needs and respond with appropriate actions[3].

Research results also indicate that the effect of teacher influence varies over a learning cycle and that the most effective known patterns of teacher influence, from the standpoint of student achievement, consist of varying levels of indirect and direct influence on student actions. Thus, highly disaggregated studies which examine the nature of teacher influence over short periods of time should be the most useful in constructing production functions designed to reveal optimal patterns of teacher behaviour[4]. Indirect teacher influence appears to be positively correlated with student attitudes, therefore differing forms of an educational production function could be expected from different objective functions.

The adolescent sub-culture has been shown to be generally detrimental to academic motivation and achievement[5]. A strong correlation between social rewards for academic excellence and the ability of high performers among high school students, regardless of other school characteristics - such as size, community, socio-economic status, expenditure per pupil - suggests that the most capable students will be high achievers only if it is socially acceptable. A significant influence upon boys' choice of an "ideal" image appears to be interscholastic athletics, perhaps because of the importance attached to these activities at a school and community level. In schools which do not have interscholastic programmes, the ideal image is more academic than in comparable schools with interscholastic athletics. For girls, the basis of the status system is more inclined to wander, though in all cases it also tends away from academic excellence.

1. Daniel Schrieber, "The Role of Universities in Supplying Help to Metropolitan School Systems", in Herbert C. Rudman and Richard L. Featherstone (Eds.), Urban Schooling, New York: Harcourt, Brace and World, 1968, pp. 33-61.

2. Ned A. Flanders, "Teacher Influence, Pupil Attitudes and Achievements: Studies in Interaction Analysis", Final Report, Cooperative Research Project No. 397, Minneapolis: University of Minnesota, 1960.

3. Ned A. Flanders, "Some Relationships Among Teacher Influence, Pupil Attitudes and Achievement", in Bruce J. Biddle (Ed.), Contemporary Research on Teacher Effectiveness, New York: Holt, Rinehart and Winston, 1964, pp. 196-231.

4. David G. Ryans has shown that pupil behaviour is more closely related to teacher behaviour in elementary grades than in secondary grades, in Characteristics of Teachers: Their Description, Comparison and Appraisal.

5. James S. Coleman, "Adolescent Sub-culture and Academic Achievement", The American Journal of Sociology, No. 65, January 1960, pp. 337-347.

Poor school environment (F) has also been shown to have a detrimental effect on measured student aptitude[1]. These results appear to contradict the findings of major empirical school characteristics studies which allege that school inputs have negligible effects on student performance.

ii) Administrative theory

Knowledge of administrative theory is less complete than is knowledge of the educational process. Furthermore, progress in forming a theory which is sufficiently well specified to be subjected to empirical estimation may be slow[2]. However, some surveys of the literature have been made which permit observations on the direction which research relating to aspects of administrative behaviour is currently assuming[3].

Studies of administrative behaviour in schools, most of which are doctoral dissertations, have focused upon the social setting from which decisions emerge. For example, studies have investigated the leadership role of principals and administrators in both the educational system and in the community[4] or surveyed the perceptions of teachers regarding the roles of principals and superintendents[5]. Few studies consider the formalization of the decision process[6]; nowhere does research attempt to identify the criteria on which policy decisions are made or the rationality and effectiveness with which these values might be pursued. One hypothesis which has considerable intuitive appeal is that decisions by school administrators are made primarily to preserve a community consensus and support staff morale. Because of the essentially conservative structure of educational systems, there is no a priori reason for expecting decisions affecting resource allocation in schools to be made with an efficiency criterion.

An understanding of the nature of and bases for administrative decisions remains a challenging problem in attempting the estimation of empirical production functions for education[7]. If the observed uses of resources in schools are to be interpreted as efficient in the sense that no greater level of student achievement could be obtained from the resources, then one must be able to presume that decision makers are aware of what constitutes best practices and, in addition, strive to adopt these policies.

1. Kenneth B. Clark, Dark Ghetto, New York: Harper and Row, 1965; Ada Hart Arlitt, "The Relation of Intelligence to Age in New Children", The Journal of Applied Psychology, VI, 1922, pp. 378-384; Helen Tomlinson, "Differences between Pre-School Negro Children and Their Older Siblings on the Stanford Binet Scales", Journal of Negro Education, XII, 1944, pp. 474-479; and Florence M. Young and Howard A. Bright, "Results of Testing 81 Negro Rural Juveniles with the Wechsler Intelligence Scale for Children", Journal of Social Psychology, XXXIX, 1954, pp. 219-226.

2. Daniel E. Griffiths, "Toward a Theory of Administrative Behavior", Administrative Behavior in Education, Ronald F. Campbell and Russell T. Gregg, (Eds.), New York: Harper and Brothers, 1957, pp. 355-356.

3. Keith Guy Hogle, Research Regarding Administrative Behavior, unpublished master's paper, University of Wisconsin, 1966.

4. Martin Gra.., A Role Analysis of the School Principalship, unpublished doctoral thesis, University of Wisconsin, 1961.

5. Arthur Blumberg and Edmund Amidon, "Teacher Perceptions of Supervisor-Teacher Interaction", Administrative Notebook, 14-1, September 1965; John Herbert Crotts, A Comparison and Analysis of the Concepts of the Role of the Elementary School Principal, unpublished doctoral thesis, University of Missouri, 1963; and William Emil Kuschman, A Study of Principal-Teacher Behavioral Perceptions based upon Selected Indices of Administrative Practice, unpublished doctoral thesis, Indiana University, 1964.

6. Two studies which do treat decision-making are Arnold Roe, An Adaptive Decision Structure for Educational Systems, unpublished doctoral thesis, University of California, Los Angeles, 1964; and Walter John Ziegler, The Bases and Processes for Decision-Making by the Superintendent of Schools, unpublished doctoral thesis, University of Southern California, 1964.

7. The need for the development of theory and its application to administrative research in all fields was noted by Herbert A. Simon in Administrative Behavior, New York: The MacMillan Company, 1958. p. 44.

SUMMARY

This part of the paper has surveyed the role of theory in empirical research into educational practices. In addition to underscoring the need for a theoretical framework for school characteristics studies, it has drawn upon the literature of educational theory and research to suggest a tentative model of the educational process. The absence of a comprehensive administrative theory has been noted as a major deficiency which must be corrected in order to supply meaningful interpretations for inter-school studies.

III

MEASUREMENT AND CONCEPTUALIZATION OF VARIABLES FOR PRODUCTION FUNCTION ESTIMATION

This part surveys the major problems that emerge in any attempt to conceptualize or measure the variables entering the educational production function. These problems occur in connection with four topics: a) the use of proxy variables to represent underlying and ill-defined variables; b) the scaling of qualitative and nominal variables; c) the reliability of observational instruments and the implications of the inter-correlations of variables for the interpretation of regression results; d) the selection of an appropriate level of aggregation both over subject areas of instruction and over grade levels.

The Use of Proxy Variables

The principal reason for the use of proxy variables in quantitative research into educational processes is the lack of a well-developed theory of the educational process. The abstractions of the educational process which are currently accepted contain variables which do not have empirically convenient counterparts. The researcher is therefore forced to resort to a variety of derived measures of the underlying concepts. Motivation and teacher quality are examples of educational system inputs for which only crude or extremely time-consuming measurement techniques exist[1]. Researchers have therefore resorted to indicant or proxy variables which appear intuitively to be highly correlated with the underlying variables. Motivation and pre-school learning have been measured by the proxy variable socio-economic class of the student's family[2]. In studies of educational production, the age of the principal structure of the school has been employed as a measure of the quality of the physical plant and of the technological vintage of school facilities[3].

The use of proxy variables introduces critical problems into both the quantitative analysis and the interpretation of results. The variable socio-economic class of family, for instance, allegedly represents the variables motivation and, in some studies, pre-school achievement. However, some sociological studies indicate that socio-economic variables have little predictive power when more basic student or school characteristics are considered[4]. It is therefore questionable whether the variation in some dependent measure attributable to socio-economic class of family is properly regarded as the result of the latent variable motivation or some unknown combination of the other variables which may be correlated with family social class.

1. One interesting but difficult measure of motivation is that of projective or thematic apperception tests developed by David C. McClelland and his associates, described in The Achievement Motive, Princeton, New Jersey: Van Nostrand, 1961.

2. Jesse Burkhead, Input and Output in Large-City High-Schools, Syracuse: Syracuse University Press, 1967, p. 43.

3. Ibid., p. 44.

4. Patricia C. Sexton, Education and Income: Inequalities in our Public Schools, New York: Viking Press, 1961; James S. Coleman, The Adolescent Society: The Social Life of the Teenagers and its Impact on Education, New York: Free Press of Glencoe, 1961; David Gottlieb, "Social Class, Achievement and the College-Going Experience", School Review, 70, Autumn 1962, pp. 273-286; and Wilbur B. Brookover, Ann Peterson and Shailer Thomas, Self Concept of Ability and School Achievement, East Lansing: Office of Research and Publications, College of Education, Michigan State University, 1962.

Proxy variables also confuse the interpretation of research results by obscuring the identity of latent variables. The easily accessible variable socio-economic class of students may measure motivation or pre-school education, as several authors have suggested; however, equally compelling arguments may be offered that socio-economic class is really a more refined and precise measure of the quality of school inputs; in studies where the decade of construction of the principal structure or the presence of science laboratories have been used to measure the quality of inputs, this argument is particularly convincing. Broad, proxy variables introduce ambiguities into research into educational production functions which are impossible to disentangle.

The Scaling of Qualitative and Nominal Variables

The objective of production function analysis is to determine the responsiveness of outputs to selected, controllable inputs; this requires that metric significance be attached to both independent and dependent variables. Economists have adopted the convenient device of resorting to money valued variables to free themselves from the task of examining the metric significance of empirical measures. The abundance of both qualitative and nominal variables in the discussion of the educational process forces one explicitly to consider questions of the unit of measurement and the point of origin of measured variables.

The problem of scaling appears in educational production function estimation in the analysis of qualitative and nominal variables. Most of the indicants of quality of school inputs and of school outputs used in studies to date are qualitative in nature. Class rank, level of motivation or ability are examples of non-metric qualitative data which are of ordinal significance only. Class rank, for instance, while providing an ordering of students, does not indicate the distance between observations: moving from 2nd to 1st in a class probably requires substantially different accomplishments and efforts than moving from 543rd to 542nd. Similarly, teachers with verbal ability measures of 140 cannot be assumed to be in some sense 40% better than those with verbal ability measures of 100. Clearly, the metric significance which can be conveniently attributed to economic variables by means of market evaluation cannot be assigned to these qualitative variables.

Nominal variables are those variables to which names have been assigned without regard for some latent dimension. Geographic region or race are examples of variables which might be assigned to arbitrary numerical categories. A less obvious example is socio-economic category based upon the profession of the father. In studies of educational production where socio-economic class is asserted to be an adequate derived measure of motivation and informal education, these nominal variables have been introduced into the regression analysis as though they were well-scaled variables. The changes in class, as persons move from categories of servant to unskilled, and managerial to professional, have been assumed in quantitative analyses to be in some respect the same.

An extensive literature on the theory of scaling exists in psychometrics[1]. While no attempt will be made to summarize techniques incorporating this theory here, some general comments should be made. The purpose of scaling theory is to provide an intuitively satisfying model of origin and interval which will adequately capture the significant dimensions of variables. Scaling is, of course, important in interpreting ordinal data; the problem of socio-economic class might, for instance, be reduced if appropriate orderings of latent variables were merged into a multi-dimensional metric which provided a semi-cardinal representation of the variable. A useful set of qualitative variables might be obtained for class which indicates the approximate distance between occupational or income groups in terms of these latent characteristics. Socio-economic class could, in such an instance, be used as an indicant for a set of underlying concepts which are rather difficult to capture or measure.

1. See, for example, W.S. Torgerson, Theory and Methods of Scaling, New York: John Wiley and Sons, 1958; P.H. Coombs, "Psychological Scaling without a Unit of Measurement", Psychological Review, No. 57, 1950, pp. 145-158; and Robert P. Abelson and John W. Tukey, "Efficient Conversion of Non-Metric Information into Metric Information", Proceedings of the American Statistical Association Meetings, 1959, pp. 226-230.

The Reliability of Observational Instruments

Standardized examinations for ability and achievement exhibit somewhat flexible standards of measurement. Any examining instrument, in addition to measuring the latent characteristic, also measures the influence of a variety of immediate factors, from the health of a subject on a particular day to the comfortableness of the testing room. Reliability coefficients reported for standardized academic tests range from approximately .85 to .98[1]. Yet, rarely are these figures quoted for sub-populations. Thus, tests may be less reliable for unusually gifted students than for poor students. To avoid serious problems of doubtful reliability, most standardized tests of ability and aptitude in the United States now report scores in stanines, which place students only in one of nine categories on a particular test. The College Board Examinations, however, still provide normalized item scores.

The laxness of the reliability of standardized tests implies that even a perfectly specified regression analysis of the educational process would be able to explain as little as 72% of measured student achievement.

The problem of reliability of observational instruments also plagues the measures of the independent variables. Several critics of major empirical studies have indicated that questionnaires employed to obtain information on student and school characteristics have not been validated. The responses obtained from elementary school children on parents' educational attainments or from teachers on their own involvement in professional development are particularly suspect.

The Choice of Aggregation Levels

The discussion of the appropriate level of aggregation over subject matter and grade levels will focus on three underlying questions. First, at what level of aggregation will the analysis produce the maximum amount of policy information per unit of research effort? Second, what are the consequences of information loss through aggregation on the policy conclusions of the analysis? Third, what are the statistical implications of aggregation? The third question will be considered in the discussion of statistical estimation below. The present section will examine the first two questions.

The level of aggregation that is appropriate to educational production function estimation depends in part upon the goals of the analysis; one must know what the policy instruments are in order to establish the proper scope of the study. The recursive structure of the educational system and the interrelatedness of endogenous and exogenous variables suggest that if one were interested in determining the optimal educational programme for a given school, information regarding input-output relationships at a fairly detailed level would be appropriate. If one knows, for instance, what the responsiveness of third grade mathematics achievement is to second grade reading skills and arithmetic achievement, then the allocations within a school programme could be improved, perhaps with dramatic results. If, however, the only form of leverage over educational practices available to policy makers is the allocation of public funds by district, then the more modest goal of determining the responsiveness of educational outcomes to expenditure per pupil is appropriate. However, even in this case, the researcher might additionally wish to identify the characteristics of those school systems which are most likely to make efficient use of additional resources.

Potential information loss through aggregation over subject matter and grade levels also raises doubts concerning production function estimation based on highly aggregated data. It is intuitively obvious that first grade reading requires fundamentally different school inputs than does advanced placement high school physics. Any educational production function that is estimated from heterogeneous data will produce a crude average function permitting only questionable interpretation. The analogue

1. Oscar Buros, Mental Measurements Yearbook, Fifth Edition, Highland Park, New Jersey: Gryphon Press, 1966.

in private sector empirical production function estimation might be to estimate a function for all manufacturing; quite obviously the resulting function would not substantially illuminate the question of the employment effects of greater investments in cotton mills.

While considerable research into the aggregation problem has been conducted, no general conclusions may be drawn[1]. The problem is essentially one of weighing the additional cost of obtaining and analyzing more variables against the loss of information or the unsatisfactoriness of more aggregative studies. The concept of aggregation bias has been developed in the context of macro-economic analysis but it does not generalize to educational production functions with ease[2]. Because one cannot at present appeal to a well-defined statistical decision theory to justify a particular level of aggregation, it remains for the investigator to evaluate the alternatives intuitively.

SUMMARY

The purpose of this part of the paper has been to identify explicitly some important problems of measurement and conceptualization of educational production function variables. The variables that one examines in studying educational production are typically derived variables that attempt to represent the qualitative significance of latent variables but which do so rather crudely; the estimated variables frequently are not cardinal measures but rather are ordinal or nominal in nature. The survey and test instruments employed in examining education are somewhat crude and only relatively reliable; a review of the reliability literature indicates that as much as 28% of the variance in output measures might be attributed to the testing instrument itself. Finally, the level of aggregation at which the analysis is conducted determines the interpretation that may be imposed upon the results of the study.

1. H. Theil, "The Aggregation Implications of Identifiable Structural Macrorelations", Econometrica, 27, 1959, pp. 14-29.
2. H.A. John Green, Aggregation in Economic Analysis, Princeton: Princeton University Press, 1964; and Walter D. Fisher, Clustering and Aggregation in Economics, Baltimore: Johns Hopkins Press, 1969.

IV

STATISTICAL TECHNIQUES FOR PRODUCTION FUNCTION ESTIMATION

This Part examines the statistical techniques that are appropriate to estimating educational production functions. In the absence of complicating statistical problems, the appropriate statistical approach to estimating a production function for education would be to obtain least squares estimates of the regression equation parameters. The regression coefficients would indicate the marginal or incremental influence of a unit change in a particular independent variable upon the output measure. These coefficients might be scaled into units of output per dollar, pound or franc of an input and hence might provide direct information on the economy with which a particular input might be employed to increase output. The t statistics and standard errors associated with each variable would assess the level of confidence that one might place in the estimated parameters. In addition, the coefficient of determination might be employed to indicate the successfulness of the analysis in explaining the educational production process.

The recursive structure of the educational process, and the problems of conceptualization and measurement of educational system variables, considerably complicate the statistical analysis. Two complicating factors will be considered here. The first is that the recursive model of education becomes subject to simultaneity bias if aggregated over time. The second is that the extensive use of proxy variables, many of which are qualitative rather than well-scaled, introduces problems both of multicollinearity and of specification of estimating equations. This section will examine these two factors of simultaneity bias and collinearity.

The problem of simultaneity bias is introduced into the statistical analysis whenever an output of the process appears during the same period of the analysis as an input. The problem of simultaneity is thus intimately related to the question of appropriate levels of aggregation over time. In the present context, simultaneity problems will emerge if the educational period being examined is lengthy enough to permit successes or failures in learning experiences to become important causal factors in determining the level of motivation and prerequisite knowledge of the student, since motivation and knowledge are important school inputs as well as outputs.

The problem of simultaneity bias may be confronted in two ways. First, one may, where data and computational costs permit, consider the educational process as a purely recursive system. A series of regressions for each stimulus-response situation may be estimated with the produced inputs being treated as exogenous independent variables in later stages of the process. Simple least squares estimation methods then produce unbiased estimates of model parameters[1]. This approach in addition generates a much more complete picture of the educational process and, hence, is much richer in policy implications. The second approach is to permit aggregation over time which will create simultaneous equation bias in the estimated parameters unless more sophisticated estimating techniques

1. H. Wold and L. Jureen, Demand Analysis, Part I, New York: John Wiley and Sons, 1952.

are used[1]. The model must be expressed in terms of simultaneous equations in which the structure of the process is carefully defined; the output variables may then be expressed as functions of the variables which are exogenous to the statistical model. The reduced form equation or equations may be used to estimate the parameters in which we are interested by a variety of statistical methods, including two-stage and three-stage least squares. The computational methods are certain to be much more complicated and the results of the analysis much less useful for policy purposes. The advantage of explicit simultaneous equation methods is that data which are not well suited to highly disaggregative recursive analysis may be employed to gain some insights into the educational process.

The problem of multi-collinearity arises when independent variables used in a statistical analysis are highly correlated. When this occurs there is no way of assuring that the causal influence unique to one variable is not assigned to another highly correlated variable; the problem is one of assigning the variation in the dependent variable, jointly explained by two or more independent variables, in a satisfactory way. This problem is particularly severe when studying educational processes because the extensive use of derived measures of latent variables results in considerable correlation of independent variables. Socio-economic class and quality of school facilities are examples of variables that usually vary together. One procedure for attacking this problem is to employ stagewise regression analysis in which the order of regression is specified a priori. The causally prior variable is entered into the regression first, and all of the variation in the dependent measure explained by the independent variable and its correlated variable is assigned to the first variable. In effect, the causal ordering of variables is used to justify assigning all of the explanatory power shared by a set of collinear variables to a first variable. When used with poorly specified models and proxy variables, this approach does little to alleviate the problem of multi-collinearity because the problem remains within the variables used in the analysis. More detailed specification of the theoretical basis for the estimation of educational production functions and direct measures of important variables should provide the researcher with insight into the pattern of interdependence among variables and suggest means of dealing with the pattern[2].

1. R. Bentzel and B. Hansen, "On Recursiveness and Interdependency in Economic Models", The Review of Economic Studies, No. 22, 1954, pp. 153-168.

2. Donald E. Farrar and Robert R. Glauber, "Multi-collinearity in Regression Analysis: The Problem Revisited", The Review of Economics and Statistics, February 1967, pp. 92-107.

V

SUMMARY

The central purpose of this paper has been to raise the most important methodological and empirical objections to empirical research into the educational production process. The large-scale, non-experimental research design employed in such well-known studies as the Coleman and Plowden Reports introduces almost insuperable problems into the analysis of the qualitative impact of controllable variables on student achievement. It has been shown that an operational theory of the educational process does not yet exist. Not only do we not have an adequate theory of feasible pedagogical alternatives, but neither do we have a theory of administrative choice by which alternatives are selected or by which we might infer the normative significance of observed allocations of educational resources. In the absence of a theoretical framework to guide in the selection of variables for analysis and in specifying functional relationships among important variants, no causal inference may be drawn from these studies. The above discussion has focused upon the validity of empirical research as a source of insights into optimal public policies towards education. The empirical research that has been conducted thus far does present interesting descriptions of many educational systems and offers some highly suggestive correlations among school and non-school variables. However, no confidence may be reasonably placed in the policy implications being drawn from such descriptive research.

The variables which have been investigated in studies of educational technology are not well adapted to rigorous quantitative analysis. This paper has suggested that the abundance of qualitative and nominal variables undermines the conclusions that have been reached. The extreme heterogeneity of educational inputs and outputs and the absence of a plausible scaling of their variables seriously cripple the statistical analysis.

The pessimistic conclusions of this paper for large-scale empirical research into educational technology suggest two general approaches to future research. First, one may accept the premise that such research is highly valuable and attempt to remedy the problems of estimating production functions for education. Second, one may conclude that less ambitious micro level research into educational technology is appropriate and endeavour to perfect less ambitious techniques and to enhance the policy value of educational research. We will briefly examine the implication of each of these viewpoints.

Application of the empirical production function concept to education demands that the theory of educational processes be improved in two important respects. First, the structure of the educational process must be better specified; current educational theory offers only very crude insights into the functional form of relationships among educational inputs and outputs. While it is possible to identify several correlates of student achievement, it is not possible to indicate whether these effects are linear, additive or independent. The argument that ability and quality of educational experience somehow interact is sufficiently appealing not to be summarily dismissed. In addition to our imprecise knowledge of the pedagogical process, we are ignorant of the determinants of the implementation of educational innovations. Several educators have noted a considerable lag between knowledge of the best educational practices and their implementation. Until systematic research is focused upon the formal decision process in education administration, this issue will remain a question.

Second, as important as the need to develop an adequate theory of education is the requirement that the theory be operational. Much of existing educational theory is essentially circular (e.g. defining teacher effectiveness by student achievement). Student ability and motivation are two important variables that are also circularly defined. More operational formulations of educational theory by educationists, as well as more objective and direct measurement of variables, are necessary before useful applications of the production function concept can be made to education.

The alternative approach to enlightening educational policy is to improve the usefulness of micro-level educational research. This alternative has several attractions. As we noted earlier, the more disaggregative are the analyses, the more policy information may be obtained from the research. Educational changes which are a priori valuable may be experimentally instituted and the consequences evaluated. The micro-educational research being suggested has been widely conducted in the form of curriculum or instructional evaluation. Several limitations in these common research strategies might be removed which would, however, increase their usefulness. First, these studies might be improved by developing a corresponding cost-effectiveness analysis; the instrumental variables under consideration might be scaled by cost in order to indicate the least cost way of augmenting educational outputs. Second, the distributional effects of educational experiences should be evaluated; the conventional evaluation strategies examine the mean achievements of experimental and control populations. A more sophisticated, multivariate statistical analysis would reveal the differential effectiveness of educational processes on students with different abilities, backgrounds or motivations. Third, the results of micro-level educational research might be integrated into a general model of education. Earlier discussions have stressed the importance of prerequisite skills in education. A variety of empirical techniques exists which might be used to examine the optimal flow of educational experiences, either in order to reduce costs or to maximize the quality and content of student achievement.

In addition to generating more abundant and useful policy information, the micro-level research strategy has the advantage of being relatively simply and inexpensive. The limited scope of the analysis reduces the need for elaborate statistical controls of intervening influences and thereby reduces the magnitude of the sample required to obtain statistical confidence in parameter estimates. The use of quasi-experimental research designs also reduces dependence on educational theory as a source of hypotheses.

This paper concludes that policy oriented research into educational technology might be most profitably pursued through refinement of relatively conventional, small-scale research designs. The grand, aggregative, production function type of study demands greater theoretical and empirical sophistication than education is presently able to provide. This paper also suggests that the policy conclusions being drawn from contemporary empiricist research into educational technology are not fully supported due to the defects of the research strategy. These problems are not merely marginal errors in data collection or interpretation but rather reflect the relatively primitive nature of our knowledge of education. These studies of education offer hypotheses and insights into education and, more important, provide indicting pictures of the equality with which educational opportunities are being distributed. The problem of securing a more efficient allocation of educational resources remains to be investigated.

Paper 2

PROBLEMS IN MAKING POLICY INFERENCES FROM THE COLEMAN REPORT

by

Glen Cain and Harold Watts[1]

1. This research has been supported by funds granted to the Institute for Research on Poverty pursuant to the provisions of the Economic Opportunity Act of 1964. The authors would like to acknowledge the helpful comments received from Dennis Aigner, Arthur Goldberger, W. Lee Hansen, Robinson Hollister, M. J. Lefcowitz, Burton Weisbrod and Walter Williams. We are most grateful to Mrs. Felicity Skidmore for editorial help.

A GENERAL INTRODUCTION

The following paper by Professors G. Cain and H. Watts of the Institute for Research on Poverty, University of Wisconsin, has been included among the papers for this Conference for several reasons. First, the document which has come to be called the Coleman Report[1] has been the subject of a major controversy among both academic analysts and governmental policy makers in the United States. As suggested in several of the papers prepared for this Conference, the Coleman Report has raised questions which are basic to the whole development of educational systems in Member countries. Thus, some discussion of the technical aspects of this controversy seems in order for this Conference. Second, the Coleman Report is significant not only for the particular controversy regarding its conclusions, but also because of the implications to be drawn from the controversy about the technical procedures used for the whole body of studies f educational processes.

Perhaps the most important lessons to be drawn from the Coleman Report controversy are those regarding the methodology of quantitative studies of education. It is in this regard that the Cain and Watts paper is particularly rich in insights. They show how the absence of a clear theoretical structure leaves the interpretation of the causal significance of many of the specific variables measured subject to a good deal of ambiguity. Further, they not only point out the limitations in the Coleman Report methodology, but they also indicate what the requirements are for an improved methodology. Third, Cain and Watts address the problems of method and inference not solely in terms of the requirements of academic rigour, as has been the case of most critics of this sort of study in the past, but more particularly in terms of the use of results in the context of policy decision-making. The points they raise with respect to policy-related inference extend beyond the Coleman Report, and even beyond studies of education in general, to the area of quantitative inference for the whole realm of public policies. For all these reasons, it was felt that participants in this Conference should have this paper made available to them.

1. James S. Coleman et al., Equality of Educational Opportunity, Washington: United States Office of Education, 1966.

I

INTRODUCTION

The aim of the Coleman Report[1] is twofold - a) to describe certain aspects of our educational system and b) to analyse the way it is related to educational achievement - with the objective of prescribing policies to change the system. In its purely descriptive aspects, the Coleman Report presents a very dismal picture of the effectiveness of our educational system in securing equal opportunities for all our citizens. Looking at educational outcomes for children from different backgrounds one finds wide discrepancies which the "American dream" has assumed capable of elimination through the public school system. These discrepancies have been authoritatively established in the Coleman Report, and the indictment and challenge they present are a crucial contribution. Although we take a critical view of this Report, nothing in our subsequent commentary can detract from the importance of the findings regarding the inequalities in the education of children of different races, ethnic groups, and socio-economic classes.

Our criticism of the Report is directed toward its analysis, mainly found in Chapter 3, in which an implicit theory of the determinants of educational achievement is posited, tested, and used to point up prescriptive policy implications. The principal theme of our discussion is that the analytical part of the Coleman Report has such serious methodological shortcomings that it offers little guidance for policy decisions. Other critics have pointed to the shortcomings that resulted from non-response to the survey and from errors in measuring certain variables[2], [3]; and the familiar uneasiness about interpreting non-experimental data has been expressed[4]. Our criticism is more fundamental in the following sense. Even if the survey data were uncontaminated by any biases from non-response, errors in measurements, and an "uncontrolled experiment", there remain the following two basic defects in the Coleman analysis.

First, the specification of the theoretical model is inadequate to support the regression analysis used in testing the model. Little or no theoretical justification is offered for the selection of explanatory variables, for their functional form, or for the inclusion or exclusion of variables under different specifications of the model. Without a theoretical framework to provide order and a rationale for the large number of variables, we have no way of interpreting the statistical results. We have no way of knowing, for example, whether a variable directly represents a policy instrument or is only indirectly related to policy control through some other unmeasured (or partially measured) relationships; or whether a variable is, indeed, supposed to be subject to policy control or is included in the model to perform a different function. (Examples of this problem are discussed below).

1. James S. Coleman et al., Equality of Educational Opportunity, op. cit.

2. Samuel S. Bowles and Henry M. Levin; "The Determinants of Scholastic Achievement - An Appraisal of Some Recent Evidence", Journal of Human Resources, 3, Winter 1968.

3. John F. Kain and Eric A. Hanushek, "On the Value of Equality of Educational Opportunity as a Guide to Policy", Discussion Paper No. 36, Program on Regional and Urban Economics, Harvard University, Cambridge, Massachusetts.

4. See the remarks of William H. Sewell, ("Review", American Sociological Review, Vol. 32, No. 3, June 1967, p. 478), of Robert Nichols, ("Schools and the Disadvantaged", Science, Vol. 154, 9th December 1966), and of Frederick Mosteller in "The Preliminary Report for Group D", 29th March 1967.

Second, in those instances where a theoretical justification for the use of a variable in the regression model is clear, the criterion used in the Coleman Report to assess or evaluate the statistical performance of the variable is inappropriate. Instead of providing information about the quantitative effect of a variable in altering educational achievement - information which would enable the reader to assess the feasibility and costliness of operating on the variable - the report provides information about a statistical measure of the variable's performance (namely, its effect on the coefficient of determination, or R^2, of the regression), which gives no clear guidance for translating the statistical findings into policy action.

The remainder of the paper is organised around the development of these points. In Part II we comment briefly on the policy objectives which determine the choice of a dependent variable - namely, a measure of educational achievement. In Part III, the core of the paper, we discuss the nature of a statistical-theoretical model necessary to handle any analysis of the determinants of educational achievement. A hypothetical and simplified example is used to indicate a relevant set of questions in terms of the objectives of social policy, and to suggest how the results from testing the statistical model should be translated into terms suitable for policy decisions. We should emphasize, however, that the example is hypothetical. The most serious gap concerning educational policy, particularly compensatory education, remains that of an inadequate theory, and we cannot fill that gap. In Part IV of the paper we do, however, discuss a few of the many specific variables which are found in the Coleman Report to at least illustrate the points made in our hypothetical example and methodological discussion.

II

POLICY OBJECTIVES UNDERLYING THE COLEMAN REPORT

A statement of a desirable or at least acceptable objective for social policy is provided by Coleman himself.

> "Schools are successful only insofar as they reduce the dependence of a child's opportunities upon his social origins. We can think of a set of conditional probabilities: the probability of being prepared for a given occupation or for a given college at the end of high school, conditional upon the child's social origins. The effectiveness of the schools consists, in part, of making the conditional probabilities less conditional - that is, less dependent upon social origins. Thus, equality of educational opportunity implies, not merely "equal" schools, but equally effective schools, whose influences will overcome the differences in starting point of children from different social groups"[1]

The task of translating the objective of equality of educational opportunity into operational terms, however, is a difficult one. The problem is twofold. First, the objective rests on a proposition - that the median levels of ability are roughly similar across racial and class groups[2] - which can be assumed but is not proven. Second, the assessment of progress towards that objective requires measuring instruments that have yet to be perfected[3].

One way to cope with the measurement problem is to rely heavily on the assumption of relative similarity in average abilities. On this basis, changes in factors (other than ability) which bring about educational achievement may be implemented, and the success of this effort may be tested by achievement scores that are correspondingly averaged over relatively large groups.

Such a focus on instruments of public policy to narrow the gaps between average levels of educational attainment across racial and economic groups has several implications:

a) The first priority is to develop a model in which the selection of variables is governed by a distinction between those variables which are amenable to policy manipulation and those which are not. The use of non-policy variables may be desirable for i) stratifying the population if we think the policy variables have different effects on different groups, and ii) controlling for intervening effects which otherwise may bias the statistical measures of the effects of policy variables. Adding non-policy variables also serves to reduce residual variation (i.e. to increase the R^2). But with the current availability of large sample sizes this may not have a high priority, particularly since

1. James S. Coleman, "Equal Schools or Equal Students?", The Public Interest, Vol. 1, Summer, 1966, p. 72.

2. The median is relatively insensitive to the location of the tails of the distribution - a fact that increases the acceptability of our working assumption. We set aside the question of how the dispersion of the distribution of innate abilities compares across groups.

3. A serious obstacle to this approach is that our current measuring instruments are clearly not able to discriminate between ability factors and achievement factors. The problem of inadequate measuring instruments is emphasized in Frederick Mosteller, "Report of the Harvard Faculty Seminar on the Equal Educational Opportunity Report, Group A", 11th May 1967, pp. 7-8, and John F. Kain and Eric A. Hanushek, op. cit., pp. 20-21.

problems of interpreting the statistical results arise as more and more variables are added, some of which inevitably overlap into the role of a policy variable.

b) A possible conflict arises between the objective of narrowing the gap between groups and the objective of raising the overall average level of each group. Certainly there would be little support for a policy which lowered average levels of performance. If, however, our prima facie evidence leads us to the assumption that the lower economic groups and disadvantaged ethnic minority groups are performing well-below their potential, then a policy which seeks to raise their performance levels may be both egalitarian and an efficient way to raise the overall average level of performance of all the groups combined. (We take up the issue of cost-effectiveness below.)

c) A similar conflict between i) reducing dispersion, and ii) raising the mean level, also exists within a group. (We should note at the outset that we must expect large variances within groups relative to that between groups. Every ethnic and economic group, after all, includes imbeciles and geniuses, stable personality types and psychotics, hard-working students and lazy students, and so on.) A strategy of compensatory education aimed at a disadvantaged group might call for raising the mean level at the expense of widening the distribution. The acceptability of this outcome would have to be examined in the particular case, but it is difficult to believe that our society is likely to undertake any policies to cope with between-group differences that will widen (or indeed severely compress) existing within-group variance.

d) It may appear trivial to suggest that the variables which serve to represent educational achievement ought to be carefully chosen and justified. The Coleman Report gathered data on several measures, but fixed on one (test scores on verbal ability) to carry almost the entire burden of the published analysis. If the several tests of achievement are measuring different "outputs", then theoretical considerations ought to dominate the choice of the most suitable "output" variable. If they are all measuring the same thing (each one imperfectly) then some, indeed almost any, linear combination of the several tests would be better than any one of them.

However, the authors seem to have postulated that one of the tests contained "it", or anyhow more of "it", and then performed the most remarkable feats of statistical augury to discover which one[1]. Perhaps other measures would have performed in the same way as the verbal ability test; we will not know until someone has tried them. But there is no indication that the choice was made on any relevant basis, and any unique properties of the measure that was used only add to the concern about the interpretation of the findings.

1. One justification for selecting verbal ability was that this variable possessed the largest relative inter-school variances. Another was that among the inter-student variances of test scores, school input variables accounted for more of the variance of verbal ability than of other test scores. It appears that what underlies these puzzling justifications is a preoccupation with "getting large R^2's" about which we will have a good deal to criticize in Part III. Suffice it to say here that the R^2 criterion is not relevant. What is relevant (but nowhere forthcoming in the Report) is a defence of such a verbal ability test as being a valid measure of educational achievement which is related, on the basis of a hypothesis concerning the determinants of educational achievement, to a specified set of school input variables. Instead, the fact that the verbal ability test is less likely to be affected by the variation of school curricula and instruction than are some of the other tests is offered as further justification for settling upon the verbal ability test ! (See: James S. Coleman, et al., Equality of Educational Opportunity, op. cit., pp. 293 ff.).

III

A SUGGESTED APPROACH TO MEASURING THE DETERMINANTS OF EDUCATIONAL ACHIEVEMENT

The following points about the analysis of specific variables as determinants of educational achievement are developed in this section. The role of a variable in affecting objectives can only take on meaning and be interpretable in the context of a carefully specified and theoretically justified model. When we have such a model in the form of a regression equation, the regression coefficient is our most useful statistic in measuring the importance of the variable for the purposes of policy action.

a) The Issue of the Significance and Importance of a Variable

In the analysis of the relation of school factors to achievement, the principal statistic offered in evidence by the Coleman Report is the per cent of variance explained. As indicated in their methodological appendix, this is because the authors are interested in assessing the "strength" of various relationships, and they believe that the per cent of variance explained provides the best general purpose indicator of "strength". It will be argued below that this measure of strength is totally inappropriate for the purposes of informing policy choices, and cannot provide relevant information for the policy maker[1].

Consider a general function expressing a relation between y and several x's, $y = f(x_1, x_2, \ldots x_k)$. What conceptual framework can be used to discuss the strength of the relation of y to, say, x_2? If we are limited to the information provided by the function $f(x_1, x_2, \ldots x_n)$, the partial derivative $\delta y / \delta x_2 = f'_2(x_1, x_2, \ldots x_n)$ is both simple and complete. In the case of linear functions, the partial derivative is a constant and expresses the change in y induced by a unit change in x_2.

It should be clear that a change in the unit of measurement will change the magnitude of such derivatives, and that any comparison among them must establish some basis for comparability among the units of measurement. In the context of an analysis of the relation of school factors to pupil achievement, it would seem evident that our interest lies in purposive manipulation of the x's in order to effect an improved performance in terms of y. We can, and should, ask for the expected change in y induced by spending some specific amount of money (or political capital, man-hours, etc.) on working a change in x_2, say, as compared with the alternative of spending the same sum on x_3. Budgetary cost is not

1. That the main purpose of the Coleman Report is to serve as a guide to policy action is made explicit and emphasized repeatedly by: James S. Coleman, "The Evaluation of Equality of Educational Opportunity", Report No. 25, The Center for the Study of the Social Organization of Schools, Johns Hopkins University, 1968.

necessarily the only basis of comparability. But unless some such basis is defined, and its relevance to policy explained, the question of "strength" has no meaning.

What basis of comparison among the x's is implied by the per cent of variance explained - which is the indicator of the "strength" of a variable used in the Coleman Report? To answer this question we will consider the common case of a linear function, the only type of function investigated in the Report.

The ordinary partial regression coefficients, b_i, for $i = 1, 2, \ldots, k$, represent the partial derivatives of y with respect to the several x's - where each x is measured in some conventional (perhaps arbitrary) unit. As indicated earlier, some adjustment of these derivatives is generally required in order to establish comparability. By using the per cent of variance uniquely explained by x_i, call it ϕ_i, as the measure of strength, the authors have implicitly assumed that x's will be rendered comparable by measuring them in units corresponding to the orthogonal (or uncorrelated) part of their respective sample variances. It is easily shown that:

$$\phi_i = b_i^2 \frac{s_{x_i}^2}{s_y^2} (1 - R_{ai}^2)$$

where the s symbol refers to the sample standard deviations and R_{ai}^2 is the coefficient of multiple determination for the "auxiliary" regression of x_i on the other (k - 1) of the x's[1]. Thus, ϕ_i represents the square of the regression coefficient which would have been obtained if:

a) each of the x's had been divided by its standard deviation discounted for its relation to other variables, and

b) y had been divided by its standard deviation, i.e.

$$\phi_i = \left(\frac{\delta_y^*}{\delta_{x_i}^*} \right)^2$$

where $y^* = \frac{y}{S_y}$ and $x_i^* = \frac{x_i}{s_{x_i} \sqrt{1 - R_{ai}^2}}$

1. What we refer to as the ϕ_i statistic is labelled the "usefulness" measure of the i^{th} variable (denoted by $\rho^2_{y, xi(p)}$) in Richard B. Darlington, ("Multiple Regression in Psychological Research and Practice", Psychological Bulletin, Vol. 69, No. 3, 1968, pp. 161-182, whose discussion of this statistic parallels much of ours and suggests several references for the interested reader. He uses:

$$\Phi_i = \rho^2_{y, x_i(p)} = b_i^2 \frac{s^2_{x_i \cdot (p)}}{s_y^2}$$

where $s^2_{x_i \cdot (p)}$ is the residual variance of x_i, i.e. the variance of x_i after controlling for all other x's in the multiple regression. His $s^2_{x_i \cdot (p)}$ is precisely equal to our $s^2_{x_i}(1 - R^2_{ai})$, and was shown in this form for the special case of a multiple regression with two predictor variables. (See equation (6) in Darlington, op. cit.). The expression, R^2_{ai}, is the same statistic as the C^2 referred to by James S. Coleman, ("Equality of Educational Opportunity", Journal of Human Resources, 3, Spring 1968, pp. 241-242) in his reply to the comment by Bowles and Levin (op. cit). Note, however, that Coleman's definition of the "unique contribution" of a variable, which involves C^2, is in error unless the variable whose contribution is being assessed has a unit variance.

It seems very difficult to find a reason why x's measured in terms of dependency-discounted-deviations, or "3-D's", are comparable for any policy purpose. Is a 3-D increment of x_1 equally costly, equally feasible, or equally appealing to the Congress as an increment of x_2? Is there, indeed, any basis for arguing that these 3-D units form a relevant set of policy alternatives such that one would have the slightest interest in how the several variables rank according to ϕ_i?

It should be clear that measuring "strength" by the usual regression coefficients, or by the Beta coefficients[1], is in general no better than using ϕ. Whether the variables are scaled conventionally or by some equally arbitrary sample-generated unit, they will usually have to be readjusted to secure comparability in the context of a specific choice problem. (This task is usually simpler if the conventional scale has not been altered, and it is more likely to be recognized as a necessary step in the analysis[2].) Although the discussion above was in terms of single variables in a given function, analogous arguments hold for groups of variables, or for the same variable in functions describing relations for different groups, regions, years, etc.

How did the choice of such an odd measure of "strength" come about? A plausible explanation is that the investigator is focussing on the "statistical significance" of the relationship. In fact the F-ratio test statistic, which is commonly used to test the hypothesis that one or several coefficients in a linear function are equal to zero, is very simply related to ϕ. When a single coefficient is tested, the F-ratio is strictly proportional to ϕ:

$$F_{1,\ t-k-1} = \frac{\phi_i\ (t-k-1)}{1 - R^2},$$

where t = sample size, and k = number of independent variables in the regression.

Where F is greater than some critical value, one commonly reports that the variable in question is significantly greater than zero at, say, the .05 level. All this means is that in order to maintain a belief that the variable in question has absolutely no effect, one must believe that the sample analysed has surmounted odds of 20 to 1 by showing such a large apparent effect. Clearly, the greater is ϕ_i or F, the greater the statistical significance and the harder it becomes for a betting man to stick to the

1. Note that:

$$\phi_i = B_i^2\ (1 - R_{ai}^2).$$

If there is only one x, i.e. k = 1, or if x_i is orthogonal to all other x's, the term involving R_{ai}^2 drops out and we have:

$$r^2 = \phi_i = B_i^2 = \left[b_i \frac{s_{x_i}}{s_y} \right]^2 = \text{the squared Beta coefficient.}$$

2. Indeed, an important advantage of the ordinary regression coefficient, b_i, is that the effect of a unit change in x_i on y is, as a matter of course, translated by the user of the statistics into terms relevant for his decision context. It has been suggested that publication of the regression coefficients produced by Coleman's research would lead to reckless and irresponsible interpretations ("Equality of Educational Opportunity: Reply to Bowles and Levin", op. cit., p. 240). This must be because either the statistics themselves or their interpreters are untrustworthy. If the problem lies with the statistics, it is hardly more responsible to publish statistics which are better behaved simply because they are definitionally limited to the positive numbers between 0 and 1, without revealing the more suspicious-looking joint products of the analysis. If the problem lies with the analysts, why give them any statistics at all?

belief that the partial derivative is zero. This is surely a very restricted and specialized meaning of "significance", since it may bear no relation to the significance (i.e. importance) a variable has for policy purposes[1].

When the regression model has included all the independent variables, the F-test (or related t-test) of the "net" or "partial" coefficients is not, of course, affected by the order of introduction of the variables into a stepwise calculation of the regression. But, the effect of a variable or set of variables (however "effect" is measured) will show up as different in the case where another set of variables is "held constant", from the case where there is no control over that other set. The only exception is when the variables to be controlled are uncorrelated with the set being examined, but this situation is present so rarely in non-experimental data that it can be dismissed[2].

When there is a legitimate interest in testing the zero-effect hypothesis, one of the variants of the F-test is available and nothing else will quite do. There is an entirely unwarranted tendency, however, to use the F-statistic (or its cousin ϕ) to indicate the more relevant kind of policy significance. To take a homely example, one might suppose that height and sugar consumption are both related to an individual's weight (among other things of course). In most contexts, height would explain more variance than sugar consumption. But to a person embarking on a weight-control programme this is not the important fact. Anyone who would seriously entertain the hypothesis that weight does not depend on height has more blind objectivity than most of us - but such a person is the only one who should care about the relative size of that test statistic. It is easy to imagine an interest in a test on the "sugar effect", but why say that it is less important or significant or strong, just because it explains less variance?

A second possible defence for the practice of evaluating variables by ϕ_i lies in its similarity to the Beta coefficient. The use of such "standardised" regression weights is usually predicated on an assumption (rarely made explicit) that the sample standard deviations used for adjusting the regression coefficients indicate a relatively fixed range of variation for the several variables. There is, in other words, some notion of "normal" limits of variation which are related somehow to the variation actually found in a population. If some x shows little variation in a representative sample drawn from an interesting population - the argument goes - then we must reduce its coefficient in order to achieve comparability with the coefficient of another x that has a larger variance.

The use of ϕ_i for comparing the effects of variables can be interpreted as the result of following this same logic farther into the labyrinth of least squares regression algebra. Specifically (as seen by the formulas on p. 76, the standardisation involved in ϕ_i is in general sensitive to the sample variances and inter-correlations for all the x's in the regression. Such a standardisation is of interest only if one feels that the entire joint distribution of regressors is both fixed in the population and well represented by the sample.

1. When ϕ is properly interpreted as a test statistic, one must keep two things in mind. a) Its relevance is limited to the zero-effect null hypothesis and b) that, as in all hypothesis tests, the power of the test is as important as the level of significance. A body of data may be unable to reject the hypothesis that some coefficient is zero, and be equally consistent with a hypothesis embodying a miraculously high effect. Alternatively, a very powerful test might reject the zero-effect hypothesis, and also reject a hypothesis that the effect is large enough to warrant any further interest in a variable.

2. An extensive controversy concerning the order of variables has appeared in the literature*. But neither critic nor defender has presented an adequate theoretical framework within which the objects of their dispute become worth arguing about.

* - Samuel S. Bowles and Henry M. Levin, "The Determinants of Scholastic Achievement - An Appraisal of Some Recent Evidence", op. cit.
- James S. Coleman, "The Evaluation of Equality of Educational Opportunity", op. cit.
- James S. Coleman, "Equality of Educational Opportunity: Reply to Bowles and Levin", op. cit.
- Marshall S. Smith, "Comments on Bowles and Levin", Journal of Human Resources, Vol. III, No. 3, Summer 1968.
- Samuel S. Bowles and Henry M. Levin, "More on Multicollinearity and the Effectiveness of Schools", Journal of Human Resources, Vol. III, No. 3, Summer 1968.
- John F. Kain and Eric A. Hanushek, "On the Value of Equality of Educational Opportunity as a Guide to Policy", op. cit.

There are many contexts, particularly in the natural processes studied in the physical sciences, when the persistence of specific sizes of the variances and correlations among some of the variables may be a warranted assumption. But it is patently absurd to postulate such invariance for variables that can be affected, directly or indirectly, by the policy alternatives that have motivated the analysis.

The use of Beta coefficients (standardised only for variance) is subject to the same sort of criticism - they retain their meaning only so long as there is no intervention by man or nature to change the variances used for standardisation. But where B_i is only crippled as a guide to policy, ϕ_i is totally disabled. The latter maintains its relevance as a description of a relationship only if we stand aside and wring our hands.

b) A Hypothetical Numerical Example

A number of the points discussed above can be grasped most readily by a review of a simple numerical example. Suppose that the relation between a suitable measure of school outcomes y, and indexes of school quality x_1 and non-school background and environment x_2, is as follows:

$$y = 1 + x_1 + 2.0x_2 + u$$

The constant term reflects an arbitrary choice of origin for the outcome measure, and we assume that x_1 and x_2 are standardised scales with zero means and unit variances[1]. The final term, u, is an unobserved disturbance term which must, in part, reflect measurement errors in y and other relevant factors such as "native ability" (whether genetic or irreversibly determined at some earlier time). This disturbance is defined to have a zero mean and to be uncorrelated with x_1 and x_2. (Assuming that x_1 and x_2 are uncorrelated with u, either singly or in a linear combination, permits us to accept the regression coefficients as unbiased measures of the effects of x_1 and x_2.) The variance of u is arbitrarily set at unity.

Now consider several alternative situations which reflect different policies with regard to the allocation of the composite bundle of factors which determine school quality, x_1. For greater simplicity we will not consider allocations that change the variation of x_1 over schools. Only the degree, and sign, of the correlation between x_1 and x_2 (ρ_{12}) will be changed. To make the policy more concrete (and more obviously hypothetical), suppose that all schools have wheels so that a fixed population of schools of various qualities can be moved around to serve an equal number of communities. A zero correlation between x_1 and x_2 ($\rho_{12} = 0$) would result from a random assignment of schools to communities. It would be changed to a positive value by moving some of the better schools from "bad" communities (as measured by x_2) to "good" ones, and vice versa. Similarly, ρ_{12} would become negative if the bad communities swapped their bad schools for good ones from the good communities. ρ_{12} would approach 1.0 if the "best" school served the "best" community, the second best school the second community and so on.

Any alteration in the way input variables are combined will change the distribution of the outcomes; for instance, a change in the variance of y is a necessary result of a change in the correlation between x_1 and x_2, given our specification of constant variances of x_1 and x_2 and constant effects (b's) of x_1 and x_2. Table 1 shows the consequences for several parameters when the correlation between x_1 and x_2 takes on several different values, ranging from 1.00 to -1.00.

1. These scalings merely simplify the numerical calculations and interpretations of the example. It could be noted that since y is not similarly standardised, there is nothing at all unconventional about a coefficient of two for the second independent variable.

Table 1. CONSEQUENCES OF VARYING CORRELATION BETWEEN REGRESSOR VARIABLES IN A SIMPLIFIED REGRESSION MODEL

Model: $y = 1.0 + 1.0x_1 + 2.0x_2 + u$

$$\sigma_{x_1}^2 = \sigma_{x_2}^2 = \sigma_u^2 = 1.0$$

$$\rho_{ux_1} = \rho_{ux_2} = 0.0$$

ROW NUMBER	PARAMETERS	I	II	III	IV	V	VI	VII
1	ρ_{12}	1.00	.90	.50	0.0	- .50	- .90	-1.00
2	ρ_{12}^2	1.00	.81	.25	0.0	.25	.81	1.00
3	σ_y^2	10.0	9.6	8.0	6.0	4.0	2.4	2.0
4	$\rho_{yx_1}^2$	.900	.712	.500	.167	0.0	.267[1]	.500[1]
5	$\rho_{yx_2}^2$	.900	.876	.782	.666	.563	.505	.500
6	$R_{y \cdot x_1x_2}^2$	.900	.896	.875	.833	.750	.583	.500
7	ϕ_1	0.0	.020	.093	.167	.187	.078	0.0
8	ϕ_2	0.0	.184	.375	.666	.750	.316	0.0
9	$R_{yx_1 \cdot x_2}^2$	0.0	.160	.429	.500	.429	.160	0.0
10	$R_{yx_2 \cdot x_1}^2$	0.0	.432	.750	.800	.750	.432	0.0
11	B_1	.312	.327	.354	.408	.500	.645	.707
12	B_2	.624	.654	.708	.816	1.000	1.29	1.114
13	b_1	1.00	1.00	1.00	1.00	1.00	1.00	1.00
14	b_2	2.00	2.00	2.00	2.00	2.00	2.00	2.00

1. The squared simple correlation coefficiencies shown here are squares of negative values for ρ_{yx_1}. All other values for ρ_{yx_1} and ρ_{yx2} in the table are positive.

In column IV one finds the simple case when x_1 and x_2 are uncorrelated - schools have been assigned to communities at random. The variance of y (σ_y^2) is equal to 6.0, and this partitions nicely into a component due to school differences with variance 1.0, another component due to community differences with variance 4.0, and a third due to the combination of factors accounted for implicitly by the disturbance term with variance 1.0. The two variables, x_1 and x_2, together account for 5/6 of the variance - 1/6 for x_1 and 2/3 for x_2 - as shown in the entries for the simple squared correlations ($\rho^2_{yx_1}$, $\rho^2_{yx_2}$) and the squared multiple correlation, $R^2_{y.x_1x_2}$.

Because x_1 and x_2 are uncorrelated (orthogonal), the incremental fraction of explained variation that is obtained when x_1, say, is added to the regression ($\phi_1 = R^2_{y.x_1x_2} - \rho^2_{yx_2}$) is equal to the fraction explained when x_1 is used alone ($\rho^2_{yx_1}$). The same is true for the increment due to x_2.

The squared partial correlations are obtained by dividing the increment due to x_1, say, by the fraction of variance left unexplained by x_2:

$$\rho^2_{yx_1 \cdot x_2} = \frac{\phi_1}{1 - \rho^2_{yx_2}}$$

$$= \frac{R^2_{y.x_1x_2} - \rho^2_{yx_2}}{1 - \rho^2_{yx_2}}.$$

The Beta coefficients, B_i, are simply the partial regression coefficients divided by the standard deviation of y, σy, and multiplied by the unitary standard deviation of x_i. The partial regression coefficients shown in the last two rows are constant, of course, because the populations have been generated by maintaining that assumption. (Columns I and VII, where x_1 and x_2 are perfectly correlated, are limiting cases - the multiple regressions would be impossible to carry out with data generated from these cases.)

The values of the various parameters listed in the columns of this table must be regarded as "population" values. A limited sample drawn at random from one of these populations could produce estimates of these parameters which would differ from the "true" values by sampling errors of the usual sort.

If the allocation of x_1 is changed from a random one by matching "good" schools with "good" communities, the correlation between x_1 and x_2 becomes positive. Moving towards the left from column IV in the table, one finds first that the variance of y gets larger. This is intuitively explained by thinking of the schools as reinforcing and intensifying the inequality found in the environments. The simple correlations shown in the fourth and fifth rows both increase as the two variables become increasingly good substitutes for each other, and the multiple correlation goes up because the constant amount of unexplained variance (from u) becomes a smaller part of the whole variance of y.

The incremental explanatory power or "unique contribution" (measured by ϕ_i) declines as ρ_{12} increases from zero, and ϕ reaches zero in the limit where $\rho_{12} = 1$. The squared partial correlations display basically the same pattern. Both are transparent consequences of the increasing interchangeability of x_1 and x_2 - as their correlation increases, having both adds very little new information. Finally, the Beta coefficients decline as a consequence of increases in variance of y. Any deeper meaning of this change must be supplied by those who have a penchant for using this scaling convention.

Consider now the consequences of allocating relatively more "good" schools to the "bad" locations and vice versa. As ρ_{12} falls from zero to negative values one finds the variance of y falling also. (See columns IV to VII.) Here the schools compensate for, or suppress, the inequality produced by unequal backgrounds.

The squared simple correlations, $\rho^2_{yx_1}$, both fall initially; $\rho^2_{yx_1}$ going to zero at $\rho_{12} = -0.5$[1]. The variance explained by x_2 falls steadily until at the limit it explains only half of the (smaller) variance of y. Beyond $\rho_{12} = -0.5$ (in columns VI and VII) the simple correlation of x_1 with y becomes negative, and in the limit it is simply a mirror-image of x_2 and thus has the same squared correlation.

The squared multiple correlation falls as the "unexplained" component of the variance becomes relatively more important. The net or unique contributions, ϕ_i, are seen to reach a peak at $\rho_{12} = -0.5$ and then to fall once more to zero as x_1 and x_2 become more identical. The squared partial correlations are seen to fall quite symmetrically on both sides of column IV where $\rho_{12} = 0$.

Finally, the smaller variance in y brings about an increase in the Beta coefficients. By this measure the effects of both x_1 and x_2 become more and more powerful; by contrast, the regression coefficients measuring their effects remain unchanged at their assigned values.

Now consider a not-entirely-hypothetical society which has shown some tendency to place its "best" schools in the "best" places and to direct its "best" efforts toward its "best" pupils. This produces an ρ_{12} somewhere between 0.5 and 0.9 - like columns II or III. An educational survey might very well find that background and environment are 4-10 times as strong as school quality if it looks at the relative size of the ϕ_i. Less extreme, but no more relevant, statements could be made by comparing the b's or B's. But what is the purpose of such comparisons? If the survey is large enough to get decent estimates of the b's, its authors could observe b_1 and infer that school quality does make a difference. It follows that moving some schools could change ρ_{12}, and shift the society's educational process toward one described by columns V or VI. Such a reallocation would substantially reduce the inequality of outcomes and attenuate the correlation of outcomes with social origins; and it would seem to be a proper sort of alternative to consider when interpreting the results of an educational survey.

It must be heavily underscored that, in terms of the model reviewed above, comparisons of the relative explanatory strengths of the two variables x_1 and x_2, whether one uses simple, partial or multiple correlation coefficients, unique contributions or regression weights, adjusted or not, a

1. Intuitively, when $\rho_{12} = -0.5$ we can think of the positive contribution of x_1 to explaining variation in y being exactly negated because of the negative correlation between x_1 and x_2. As the negative correlation between x_1 and x_2 gets larger in absolute value than -0.5, the true positive effect of x_1 is more than offset in the simple relation between x_1 and y (when x_2 is not held constant).

pointless. If one is concerned with assessing the possible effects of educational policy, comparisons of any kind with the effect of "control" (i. e. non-policy) variables are silly. Moreover, all the statistics involved in the comparisons, except for the unadjusted regression coefficients, are dependent upon the particular policies pursued when the data were collected. Their use runs the risk of declaring a policy feeble simply because historically it was not vigorously applied.

In the example shown in Table 1 the "best" allocation to achieve equality calls for a perfect negative correlation between x_1 and x_2. By this allocation the variance of y is reduced to a minimum (=2). It should be noted that educational policy might also change the mean and/or the variance of x_1. With these added degrees of freedom it would be possible, in principle, to eradicate all gross association of y with x_2, and - as an added option - reduce the variance of y to the absolute minimum introduced by the unobservable variable u.

c) The Need for a Theoretically Justified Model Relevant to the Policy Context

In general terms one may view the Equality of Educational Opportunity Survey as providing information on the joint distribution of a large number of variables. The analytical effort should be directed toward answering questions about how new or altered policies (more particularly educational policies) would change various characteristics of that joint distribution either directly or indirectly. To do this, one must have a consistent and complete set of specifications concerning: i) which characteristics of the joint distribution are constant, ii) which can be changed directly by specific activities (policies), and iii) which ones must therefore be determined by the assumed structure and prescribed policy.

This set of specifications is commonly termed a theory or model. In the Coleman Report there is no explicit discussion of a consistent theory of this sort. Some theory, of course, must underlie any sort of policy prescription. It is not that one can choose to draw conclusions from the objective facts alone, without the aid of any theory, but that if one leaves the theory implicit, ambiguous and obscure, contradictory possibly nonsensical or even self-contradictory premises go unnoticed.

The theoretical structure of the simple model discussed above asserts that the functional relation between y and x_1, x_2, and u, can be approximated satisfactorily by a linear and additive function, with coefficients that would remain fixed under policies designed to change the distribution of x_1 and/or x_2. Similarly it is assumed that the mean and the variance of the disturbance variable, u, will be unaffected by policies aimed at affecting y via x_1 or x_2. The objective of policy is taken to be some optimal combination of high average level of outcomes (mean of y), minimal inequality (variance of y) - at least as the variance or inequality is affected by inter-group differences - and easy class mobility (minimal correlation of y and x_2).

The tools of educational policy are taken to be measures that would shift the mean of x_1, compress or expand its variability, and/or revise the correlation between x_1 and x_2. If one wishes to consider social policy more broadly, similar alternatives for changing the distribution of x_2 would be available. Within the structure so far specified it is possible to deduce the effects on the marginal and conditional distribution of y for any particular change in the x_1 or x_2 distributions. If no further restriction relevant information is added, any particular goal in terms of the basic objective can clearly be achieved by a wide range of different manipulations of the x_1 and x_2 distributions. The question of relative strength, in the sense of ability to manipulate y, can now be seen to be meaningless - remembering that the scaling of x_1 and x_2 was arbitrary to begin with. Each of them can be used to achieve the objective so long as unlimited freedom is available for changing the mean, variance and correlation. If x_2 is not manipulable by educational policy, on the other hand, who cares how effective it might be if it were?

Consider, however, a very simplified situation in which the objective is to close a substantial gap between the mean value of y for negroes and the mean for whites. Assume that the function above holds for negroes, and that one's policy choices are limited to changing - at most - the mean value of x_1 and x_2 for negroes. Which policy or combination of them one chooses will depend on further information about the costs of each alternative. Costs may be in terms of dollars, time, political consensus or all three - but must be made explicit.

Indeed, one might, for purposes of policy analysis, scale the variables available for manipulation so that a unit change in x_1 is an equally costly (or time-consuming or consensus-using) alternative to a change in x_2. If an "Iso-chunk" of x_1 is defined to be a $1 billion worth, each one must be a fifth as large as the original unit costing $5 billion - hence its coefficient must be 0.2 (i.e. the old $b_1 = 1$ coefficient multiplied by its new unit of measure, 0.2). Similarly, an "Iso-chunk" of x_2 is only four per cent of an original unit priced at $125 billion, and hence its coefficient must be 0.08.

Several variations on the "Iso-chunk" idea can be specified. Take as given the relation between "output", y, and "inputs", x_1 and x_2:

$$(1) \quad y = a + b_1x_1 + b_2x_2 + u$$

Suppose first that the "costs" of alternative mixes of x_1 and x_2, in terms of any scarce item one finds important, are given by:

$$(2) \quad C = c_1x_1 + c_2x_2$$

One may now rewrite equation (1) in terms of "Iso-chunks" which correspond to the amount of x_i arrived at by using one unit of whatever "cost" consists of - dollars, man-hours, class-hours:

$$x'_1 = c_1x_1$$

and

$$x'_2 = c_2x_2$$

Thus, "Iso-chunks" (read dollars or hours) of C spent in changing x_i can be substituted in (1) for the x_i:

$$(3) \quad y = a + B_1x'_1 + B_2x'_2 + u$$

$$\text{where } B_i = \frac{b_i}{c_i}$$

We may call these B_i "bet coefficients" - derived from Israeli pronunciation of the Hebrew name for the corresponding alphabetic character[1].

The bet coefficients give quite direct answers as to which use of the scarce item C yields the largest increment in y. To the extent that relations (1) and (2) adequately reflect the way the world works, one

1. Professor Arthur S. Goldberger coined this felicitous terminology.

could confidently proceed to add to the existing educational process by directing all available C into the x_i for which B_i is the largest.

Unfortunately, one does not usually have that much confidence in a couple of simple linear relations. Commonly, relation (1) will be estimated on the basis of a limited sample, and one's confidence in extrapolations beyond the range of observed combinations of x_1 and x_2 deteriorates rapidly[1]. Moreover, one would rarely encounter a "cost function" as simple as the one in (2) - usually there will be diminishing returns causing marginal costs to rise beyond some point. Bet coefficients derived as above ought, therefore, to be interpreted as reflecting, at best, the relative effectiveness of variables in that vicinity of the data over which a linear approximation is deemed to be "sufficiently accurate", taking into account reservations about both relation (1) and relation (2).

1. This information on the reliability of the estimate is given by the confidence interval computed for the bet or regression coefficient. Our emphasis on the expected value of the B_i (or b_i) does not imply that we believe a decision maker would have no interest in the confidence interval. Indeed, one can imagine cases when a decision maker has some asymmetric subjective utility weighting scheme such that zero or negative values would be deemed so critical - more than offsetting the equally probable high positive values - that a B_1 which was (slightly?) lower than a B_2 would still be selected if the confidence interval of B_1 were sufficiently tighter. Such cases ought to be explicitly argued, however. This proper usage of the confidence interval does not warrant using the ϕ_i statistics, instead of the bet coefficient and classical inferences about it, as the primary criterion of policy choices.

IV

INTERPRETING SPECIFIC VARIABLES IN THE COLEMAN REPORT

The absence of any explicit theory of educational achievement is the chief source of the difficulty in interpreting the statistical results of the Coleman Report. We can illustrate the problem by discussing some of the variables used in the Report.

a) Attitudinal Characteristics of the Student

One remarkable finding of the Report's analysis is the high partial correlation of fate control/personal efficacy variables with the verbal ability score used as a measure of educational outcomes[1]. The relation was particularly strong (by the Report's criterion) among minority group children. Without a theory, however, we cannot answer the following types of question:

1) Is this variable itself merely a reflection of (perhaps "caused by") educational achievement? One can easily imagine situations in which educational accomplishment would instill confidence in a youngster and produce a high score on the measure of this variable[2].

2) Is this variable important only because it is related to various objective factors about the student's family, community, and school environments, which are not fully measured in the model, and which "really" explain both school performance and the fatalism score? This set of relations would again be quite plausible on a priori grounds[3].

   Under situations (1) and (2) above, we can say no more than the following. Either changes in the variable, "control over one's fate", are unattainable unless performance on the other objective variables is changed; or, if some change in the score could be induced (by, say, counselling), there is no reason to believe educational performance would change.

3) What if, contrary to (1) and (2), the fatalism variable is a personality trait that does have a separate influence on educational achievement? We still need to know how policy can change the trait to make use of our finding. Clearly these attitudes may be quite congruent with an objective assessment of the situation children find themselves in. If so, the school may be severely limited in its ability to reorient such attitudes (one may have to reintroduce prayer).

1. A number of questions in the survey attempted to measure the student's sense of control over his environment and his sense of fatalism.

2. Precisely this specification of the causal relationship is put forward in: Alan Wilson, "Educational Consequences of Segregation in a California Community", in Racial Isolation in Public Schools, United States Commission on Civil Rights, Washington, GPO, 1967, Vol. II, pp. 192 and 206.

3. The report explicitly notes that the simple correlations of verbal ability and the fate control variable are similar to the inter-correlation among the achievement variables. A finding which seems consistent with the interpretation that these attitudinal variables are just another means of measuring the joint output of school and non-school processes impinging on a child's development*.

* James S. Coleman et al., Equality of Educational Opportunity, op. cit., p. 319.

A verdict of helplessness may have to be passed on the schools. But the evidence in the report supports it neither by adding to our knowledge of the causal relation nor by indicating a low pay-off from interventions within that relation.

b) Characteristics of the Student's Peer Group

In a review of the Report's findings, Harry C. Bredemeier notes: "More important than all school characteristics and teacher quality for negro students is the degree to which the other students in their schools have the following characteristics: their families own encyclopedias, they do not transfer much, their attendance is regular, they plan to go to college, and they spend rather much time on homework"[1] He notes in a footnote, "I assume no one will infer from this that the solution is to put encyclopedias in everyone's home".

But, is such an inference less satisfactory than making no inference? Is it any more naive than the presentation of the vague theoretical framework that permits us almost no grounds for saying how we should interpret the "significant positive coefficient" of the encyclopedia variable? Consider the following interpretation:

Encyclopedia ownership is a variable that indicates an intellectual atmosphere in the home conducive to schooling, and/or a measure of affluence that is not fully captured in other measures (of affluence) in the model, and/or a measure of parental attention or affection that contributes to the student's emotional stability and, thereby, to school performance - any or all of which factors create the positive peer group influence.

Presumably, this interpretation is "more sophisticated" than the inference Bredemeier noted. But is it more helpful? Indeed, what our hypothetical theory has told us up to now is that: i) if it is intellectual atmosphere that underlies the relation, the variable has probably no policy significance since we do not know much about changing intellectual atmosphere. If we thought we did know something about how to make the change, we would need to know the specification of the relation between encyclopedias and intellectual atmosphere. ii) If it is affluence that underlies the relation, then we need to ask our theory to translate a unit of encyclopedias to a unit of wealth (or income flow) so that we know how much of a change in income will be necessary to yield the changes in educational performance.

We could continue these "if" questions almost indefinitely; but let us summarize the function of our hypothetical theory by saying that it has forced us to consider the possible tortured interpretations we have to make, or preposterous policy actions we might have to follow, as a consequence of such cavalier inclusion of ad hoc variables in our model.

c) Environmental Characteristics

The Coleman Report stressed that the influence of the regional and urban location of the school and the socio-economic status of the student body in the school were highly important in explaining a student's educational achievement. A theoretical proposition underlying the authors' interpretation of this finding was that the environment is exogenous and "causally prior" to such factors as school resources; so that an appropriate procedure was to enter the former variables, note the contribution to R^2, and then add the school resource variables and observe their additional contribution to R^2. Other demurrers to this procedure, quite apart from the issue of the R^2 criterion, may be mentioned.

If families select their residence on the basis of the quality of school, residence is neither exogenous to the process nor causally prior to the school resources variable. Particularly with regard to the racial composition of the school, the phenomenon of selective migration may be confounding the results. For

1. Harry C. Bredemeier, "Schools and Student Growth", The Urban Review, April 1966, p. 29.

example, if a large percentage of whites in a school or a large percentage of high socio-economic groups appear to have a positive effect on the educational performances of negroes or low SES groups, we should consider the hypothesis that the latter families have strong "tastes" for a high quality education for their children and have moved to a district where the school has a favourable reputation. The observed positive effect of the environment on the educational achievement of disadvantaged groups may therefore be overstated, since some of the effect stems from the unmeasured personal traits of the families; it is further possible that some effect is attributable to the beneficial resources of the school[1].

What theory of educational achievement justified "urbanness", "Southernness", etc., as causal factors - except insofar as these traits are related to such specific variables as the family characteristics and quality of schools found in these areas. There is a real danger that such location variables serve only to attenuate the influence of other variables, of interest when such other variables are unmeasured or measured with a large error component.

d) Teacher Quality

One type of variable that belongs in the category of school resources over which we have some degree of policy control is "teacher quality" - itself a composite concept made up of several variables. The conclusion in the Report about teacher quality appears to strike a rare optimistic note regarding the beneficial influence school resources can have in compensatory educational efforts. The Report states on page 317 that "a given investment in upgrading teacher quality will have the most effect on achievement in underprivileged areas". Surely, the theoretical justification for this variable should be quite firm. Moreover, the wording of the Report's conclusion exactly fits the criterion we have requested for assessing each variable.

Unfortunately, the statistical evidence in support of the finding the authors present concerns "variance explained". "Given the fact that no school factors (excluding student body composition) account for much variation in achievement, teachers' characteristics account for more than any other." And, "by the 12th grade, teacher variables account for more than nine per cent of the variance among negro students, two per cent among white students" (page 325). It is perhaps superfluous to mention again that this ranking of importance of a variable in terms of variance explained does not tell us what the "bet coefficients" are, nor permit us to derive them; therefore, the conclusion about a "given investment in upgrading teacher quality" for underprivileged areas is not supported. If, for example, the variance of verbal ability were large among teachers of negro students and the educational achievement scores had a relatively small variance, the high partial correlation coefficient (and β) of this variable would be consistent with a small value for the bet coefficient - even setting aside cost considerations. (See the formulas on pp. 76 and 84 of this paper).

e) School Resources

Perhaps the single category of variables most susceptible to policy manipulation is that of school resources. Unfortunately, the variables used to measure school resources are very much like the "encyclopedias in the home" we discussed above. It is difficult to know whether, for example, library books or laboratories are supposed to represent their own effects, per se, or whether they are supposed to represent a more extensive collection of items under the rubric of school facilities (or some other concept of school characteristics).

One can argue for either interpretation. On the reasonable assumption that libraries and laboratories are, and would be, closely linked to an underlying specification of the usage of these facilities,

1. The possible misallocation of the effect is more likely if the student family characteristics or the school resources variables are measured with considerable error. That a good deal of error is present in the measure of these variables has been strongly argued by Bowles and Levin (op. cit.) and by Kain and Hanushek (op. cit.).

we could treat libraries and laboratories as proxies for the "usage" concepts, which in turn can be plausibly linked to educational performance. Given this, the reader might further surmise that the two variables must be standing solely for their own effects, for otherwise the authors would have included the other items.

If, on the other hand, it is naive to assume that facilities present are facilities used, and if it would have been overly burdensome to include all relevant items in the survey, then we can more readily accept the argument that the included variables are meant to be representative of some different and/or larger collection. If so, we need to ask: a) what are these other variables, and b) what is the specification (i.e. regression equation) by which they are linked to the other variables. This really breaks up into two other questions: how accurate is the representation (i.e. how strongly are they correlated), and what is the quantitative magnitude of the relation (i.e. what are the regression coefficients linking the full set of variables to the proxy variables[1])?

The sort of questions we have been posing serve to illustrate the analytical weaknesses noted above. If the questions we have raised are overly demanding of the state of theoretical knowledge about the educational processes, we can only ask that this shaky base be made explicit. Perhaps researchers will be led to work with a more simplified model that can be well specified and interpreted - better this than a complex model that defies interpretation.

1. The complexity of this specification need not be exaggerated. There are many decision contexts in which proxy variables may represent a bundle of heterogeneous components, and it may not be worthwhile or expedient for the decision-maker to distinguish among the components to determine their separate measures of effectiveness. What is necessary, however, is some translation of a unit of the proxy variable into a unit of the larger bundle (along with, eventually, some measure of the costs of the larger bundle).

V

CONCLUSION

We are aware that a demand for theoretical rigour may be likened to a request for virtue. But we hope that the discussion in Parts III and IV has been sufficiently specific so that both the interpretation of the Coleman Report and the design of further studies will be improved.

Our criticism of the Coleman Report has been aimed at its methods and not at its substantive findings. The questions we have raised about the statistical and methodological techniques in the Report should be viewed as reinforcing the challenge to the "educational establishment"[1] to provide evidence on the effectiveness of their programmes, especially compensatory educational programmes. Nor should any research into the determinants of educational achievement overlook the potential contribution that may stem, however indirectly, from the simple improvement in economic status of the student or his family or the families of his fellow students.

1. The term was used by Daniel P. Moynihan in the context of his criticism that "educationists" - administrators, teachers, research personnel - have shirked their responsibilities to evaluate their performance and have attempted to use "technical" criticism of the Coleman Report as an excuse for continued inaction: (Daniel P. Moynihan, "Sources of Resistance to the Coleman Report", Harvard Education Review, Vol. 38, No. 1, Winter, 1968, pp. 23-36).

Paper 3

EDUCATION AND INCOME: A STUDY OF CROSS-SECTIONS AND COHORTS

by

R. Hollister[1]

1. Mrs. Nancy Williamson and Paul Christianson assisted with research, and Harold Watts and Glen Cain provided guidance at several stages in the writing of this paper. I also benefited from discussions with Tom Ribich, Bernie Saffran, and Patrick Shima. None of these persons should be regarded in any way as responsible for any errors which may be found in this paper.

INTRODUCTION

The rapid growth in the application of economic analysis to the role of education in society during the past decade has been largely due to the development of the concept of human capital, a concept admirably expounded by Schultz[1], Becker[2], Mincer[3] and others. It may be argued that the wide acceptance of this conceptual approach to educational problems owed no little to the fact that its advocates were able to show its empirical relevance, for the difficulty in applying investment concepts to education is that, since the return on education accrues during the whole working life of the recipient, empirical tests of the human capital concept would normally require a lifetime of data gathering. This seeming barrier to empirical testing was ingeniously circumvented by making use of cross-section profiles of incomes as related to education and age, instead of the unavailable time-series. It was thus possible to buttress the theoretical concepts with convincing empirical evidence.

Though time-series on the relationship between income, education and age are still scarce, there are now enough successive cross-sections available to make possible at least a rough examination of the income patterns actually experienced over ten to twenty years by some cohorts in the United States. This paper is an attempt at such an examination.

It should be said at once that the comparability of the data over such long periods of time raises virtually insurmountable problems, and no great effort has been made here to ensure comparability. There must, therefore, be some doubt about the extent to which features in analysis reflect meaningful social and economic forces at work; in fact, they may reflect no more than the incomparability of data. The discussion of data problems is relegated to the footnotes of the tables.

The issues dealt with here have been touched upon briefly by several authors: Becker[4], Miller[5], Ben-Porath[6], Lansing and Sonquist[7], Griliches[8]. Our analysis, however crude, goes somewhat further in that it deals with the specific issue of the differences between the cross-section profiles of age-education-income and the time-series experience of cohorts.

1. Schultz, Theodore W., "Investment in Human Capital", American Economic Review, March, 1961.
2. Becker, Gary S., Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, New York: National Bureau of Economic Research, 1964.
3. Mincer, Jacob, "Investment in Human Capital and Personal Income Distribution", Journal of Political Economy, August, 1958.
4. Becker, Gary S., op. cit.
5. Miller, Herman P., "Lifetime Income and Economic Growth", American Economic Review, September, 1966.
6. Ben-Porath, Yoram, "Lifetime Income and Economic Growth: Comment", American Economic Review, September, 1965.
7. Lansing, John and Sonquist, John, "A Cohort Analysis of Changes in the Distribution of Wealth", I.L. Soltow (ed.), Six Papers on the Size Distribution of Wealth and Income, Studies in Income and Wealth, No. 33, National Bureau of Economic Research.
8. Griliches, Zvi, "Notes on the Role of Education in Production Functions and Growth Accounting", NBER Conference on Research on Income and Wealth, Madison, Wisconsin, November, 1968.

I

THE DECISION ON EDUCATIONAL INVESTMENT

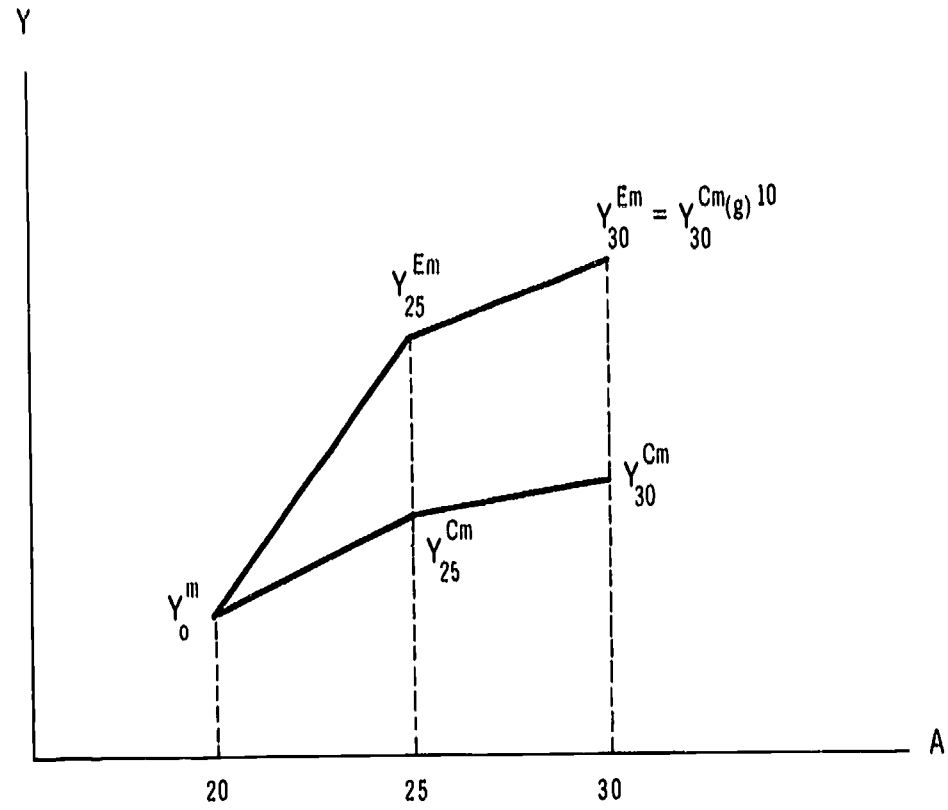
Before embarking on the empirical work, let us examine the framework within which it is to be set, and which is implicit in past studies of the rate of return on human capital using cross-section profiles of age-education-income. The following description, though distorted, has at least the merit of simplicity.

At a given point in his lifetime, say, at the age of 13 or 17, the individual is faced with a decision: he can leave school and enter the labour market, or he can continue his studies. In the latter case, he must incur at least the cost of foregoing the wages he could have earned if he had left school, and perhaps that of expenses involved in further schooling. In terms of human capital theory, he must consider whether investment in further education (through the cost incurred) will increase his future earnings to such an extent that the present value of these additional earnings exceeds the present value of returns he could obtain by investing the same amount (i.e. the cost incurred) in some other way; in other words, whether the rate of return on further education exceeds the rate of return on alternative investments open to him.

In order to make such a decision, the individual must be able to estimate, if only roughly, the increase in earnings which he may expect from additional years of education. It is suggested that such an estimate can be arrived at by examining age-education-income profiles of people currently in the labour force. For example, by taking the difference in typical (median or mean) earnings at given ages of those with 12 years of education and those with 8 years, one can trace the likely age pattern of differences in earnings due to additional years of education. By applying the appropriate discount factor to the earnings differential at each age and totalling, one can determine the present value of increased earnings likely to accrue during the individual's working life as a result of the additional education. However, one further adjustment needs to be made: earnings, in general, increase over time, and this must be taken into account. Thus, a person with a given level of education might expect, by the time he will have reached the age of, say, 50, to be earning more than the present 50-year-old with the same level of education. He will earn more as a result of the economic growth which will have taken place during the period in question. Thus, the cross-section earnings differential for each age due to additional education should be multiplied by a factor reflecting the compound annual rate of growth of earnings due to economic growth for the period in question. One can take the long-term rate of annual growth over the recent past as the expected rate of annual growth.

The way in which the expected income chart is constructed is illustrated in Diagram I. A person aged 20 in year n, with a level of education m, would observe the incomes in that same year n of those older than himself, but with the same level of education m, (Y^{Cm}_{25}, Y^{Cm}_{30} in the diagram), and to these he would apply the value of the annual rate of long-term growth, g, compounded for the number of years it will take him to reach the given age. Thus, his expected income for age 30 would be $Y^{Em}_{30} = Y^{Cm}_{30} (g)^{10}$.

Diagram I



Now, supposing that an individual had in fact approached the decision with respect to education in this fashion, how would his actual earnings experience have compared with the above prediction? In other words, do the shifts in the cross-section profiles due to general economic growth operate in a neutral fashion (we draw this terminology from Ben-Porath[1]) with respect to age and educational levels? If these shifts are neutral, then the present value of additional years of education for different levels of education, as actually experienced by age cohorts, will be in the same proportions as the estimates derived from the cross-sections.

While this rather simple framework will be the basic one used in the examination of the data, there are some more complex issues, often discussed in the context of human capital theory, which should be noted before we turn to the data. Basically, the posing of the educational investment decision problem in terms of an income stream constructed from cross-section observations and multiplied by a compound growth factor (by which income stream is usually reduced to an equivalent present value) is a convenient way of summarizing a more intricate set of forces relevant to the problem. In reality, one uses the internal rate of return to express in a single value all the complex supply and demand factors which might be expected to operate during the working life of the particular cohort of individuals whose education decisions are being made at a given time: a complex set of simultaneous supply and demand functions have been reduced to a single present value or rate of return. In estimating an expected present value of an income stream from cross-section data in the fashion described above, one avoids the identification of particular supply and demand shifts which are likely to interact to produce an expected lifetime rate of

1. Ben-Porath, Yoram, "Lifetime Income and Economic Growth: Comment", p. 827, op. cit.

return on educational investment. Broadly, the rates of return on present values of investment in education for a particular cohort of individuals coulb de affected by: shifts in relative demand for factors of production resulting from changes in the composition of demand for final goods or from technological changes in production relationships; shifts in the relative supply of different types of educated labour due to the educational investment in that cohort and in succeeding cohorts who will compete in the labour market during their working life; shifts in relative earnings associated with any asymmetrical effects of business cycle fluctuations on types of workers. With these complex interacting factors to be taken into account, one can appreciate the appeal of the simplification afforded by concentrating on the cross-section present value or rate of return estimates. If, in fact, shifts in the cross-section age-education-income profiles due to economic growth are neutral, it is not necessary to attempt to untangle the various forces at work. On the other hand, given this array of interacting forces, our traditional theory would lead us to be rather surprised (see Griliches[1]) if such neutral shifts were observed. No mechanism generating interactions such as to result in this type of neutral shift has yet been specified. In fact, this is an inherent limitation of the analysis which follows - we have not specified a satisfactory model of supply and demand interactions to be tested against these data. The best we can do for the time being is to use the simple framework outlined above to examine the extent of neutrality of shifts and to attempt to associate, in a rather haphazard fashion, any deviations from neutral shifts with some of the demand, supply, or cyclical factors.

1. Griliches, Zvi, "Notes on the Role of Education in Production Functions and Growth Accounting", op. cit.

II

THE 1939 TO 1959 EXPERIENCE

By combining the data from the censuses of 1940, 1950, and 1960, it is possible to get three "snapshots" of twenty years of work experience for various age cohorts (problems of comparability of the various years are noted in notes 1-4 to Table I). It is therefore possible with these data to make the sort of approximate comparison suggested earlier; twenty-year segments of the actual experience of various cohorts can be compared with what would have been predicted if expected income had been estimated on the basis of 1939 cross-section profiles of income by age and education level (with an expected growth rate applied as outlined above). It will also be possible to check the neutrality of shift between 1939 and 1949 and 1959 in the cross-section profiles.

In order to make this comparison, it is necessary to choose an estimate of the growth in incomes due to general economic growth. Following Becker[1] and (quoting) Cain[2], we note that: "The growth rate of the rise of real wages is basically composed of a) quality constant labor (due in large part to a slower growth in the supply of labor than the demand for labor) and b) the increase in the quality of labor. We want to allow for the secular growth stemming from a)". Therefore, we take the rate of growth of output per man-hour from 1925 to 1960 of 2.5%. Now, in order to remove the portion of this growth rate due to b) above, we reduce this rate by 25%. The resultant rate of secular growth in incomes is 1-7/8% per annum.

The expected income path for the group of individuals of a given age and education in 1939, for example those aged 20[3] who completed 12 years of education, is formed by applying the value of the compound growth rate to the 1939 median income of those with 12 years of education aged 30, 40, 50, etc. The compound growth rate is $(1.01875)^9 = 1.1819$ for age 30, $(1.01875)^9 = 1.4231$ for age 40, etc. The resultant expected income path as of 1939 is shown in Chart I. We can proceed in a similar fashion for each age-education cohort, constructing an expected income path as of 1939 on the basis of the 1939 cross-section and the expected secular growth rate.

The expected income paths thus constructed, we can make a comparison with the actual income path over the period 1939-1959 by tracing the actual observed income of the cohort in the 1949 and 1959 cross-section data. For example, those aged 15-24 in 1939 were born in 1915-1924, and for this group's actual income experience we take the observed income of those aged 25-34 in 1949, and 35-44 in 1959. This juxtaposition of the expected income path based on the 1939 cross-section and the actual income path for 20 years provides the type of contrast discussed earlier. Comparisons of expected and actual income paths for various age cohorts are provided in Charts IIa-i.

1. Becker, Gary S., Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, p. 73, op. cit.

2. Cain, Glen, "Benefit/Cost Estimates for Job Corps", Discussion Paper No. 9-68, Institute for Research on Poverty, University of Wisconsin, Madison, Wisconsin, 1968, p. 42.

3. Charts and tables below the data are for 10-year age groupings, except for the first age group centred on 21. This initial group is 18-24, but subsequent groups are 10 years, e.g. 25-34, 35-44, etc.

Table I. PRESENT VALUES - EXPECTED AND ACTUAL[1,2,3]
CALCULATED AT A DISCOUNT RATE OF 6%

Table Ia.
Cohort Age Born 1895-1904. White Males

	(1) PV 16+	(2) PV[4] 16+ Becker	(3) PV 12	(4) PV 8	(5) (1)-(3)	(6) (2)-(3)	(7) (3)-(4)
a) Actual	35,176.38	35,716.99	23,186.64	17,765.44	11,989.74	12,530.35	5,421.20
b) Expected	33,109.06	35,625.67	23,813.25	15,211.26	9,295.81	11,812.42	8,601.99
Ratio Actual / Expected.					1.29	1.06	0.63

Table Ib.
Cohort Age Born 1895-1904. Black Males

	(1) PV 16+	(2) PV[4] 16+ Becker	(3) PV 12	(4) PV 8	(5) (1)-(3)	(6) (2)-(3)	(7) (3)-(4)
a) Actual	16,152.79	n.a.	12,012.76	10,048.69	4,140.03	n.a.	1,964.07
b) Expected	13,406.36	n.a.	9,922.64	8,014.68	3,483.72	n.a.	1,907.96
Ratio Actual / Expected.					1.19		1.03

Table Ic.
Cohort Age Born 1895-1904. Total Male Population

	(1) PV 16+	(2) PV[4] 16+ Becker	(3) PV 12	(4) PV 8	(5) (1)-(3)	(6) (2)-(3)	(7) (3)-(4)
a) Actual	34,180.97	34,709.75	22,526.62	17,004.75	11,654.35	12,183.13	5,521.87
b) Expected	31,898.64	34,327.69	23,080.85	14,516.54	8,817.79	11,246.84	8,564.31
Ratio Actual / Expected.					1.32	1.08	0.64

Table I. (continued) PRESENT VALUES - EXPECTED AND ACTUAL[1,2,3]
CALCULATED AT A DISCOUNT RATE OF 6%

Table Id.
Cohort Age Born 1905-1914. White Males

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	PV 13+	PV[4] 16+ Becker	PV 12	PV 8	(1)-(3)	(2)-(3)	(3)-(4)
a) Actual	63,187.57	63,508.02	41,603.75	32,753.77	21,583.82	21,904.27	8,849.98
b) Expected	61,720.57	65,940.43	44,176.35	29,537.45	17,544.22	21,764.08	14,638.90
Ratio Actual / Expected					1.23	1.01	0.60

Table Ie.
Cohort Age Born 1905-1914. Black Males

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	PV 16+	PV[4] 16+ Becker	PV 12	PV 8	(1)-(3)	(2)-(3)	(3)-(4)
a) Actual	33,428.02	n.a.	25,104.57	20,632.62	8,323.45	n.a.	4,471.95
b) Expected	26,953.17	n.a.	20,276.61	16,214.85	6,677.16	n.a.	4,061.76
Ratio Actual / Expected					1.25		1.10

Table If.
Cohort Age Born 1905-1914. Total Male Population

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	PV 16+	PV[4] 16+ Becker	PV 12	PV 8	(1)-(3)	(2)-(3)	(3)-(4)
a) Actual	61,742.47	62,055.41	40,744.52	31,682.15	20,997.95	21,310.89	9,062.37
b) Expected	60,280.06	64,398.57	43,123.68	28,206.08	17,156.38	21,274.89	14,916.80
Ratio Actual / Expected					1.22	1.00	0.61

Table I. (continued) PRESENT VALUES - EXPECTED AND ACTUAL[1,2,3]
CALCULATED AT A DISCOUNT RATE OF 6%

Table Ig.
Cohort Age Born 1915-1924. White Males

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	PV 16+	PV[4] 16+ Becker	PV 12	PV 8	(1)-(3)	(2)-(3)	(3)-(4)
a) Actual	82,566.95	82,566.95	63,875.02	49,304.28	18,691.93	18,691.93	14,570.74
b) Expected	80,368.18	84,044.34	58,727.44	39,776.65	21,640.74	25,316.90	18,950.79
Ratio $\frac{\text{Actual}}{\text{Expected}}$.					0.86	0.74	0.77

Table Ih.
Cohort Age Born 1915-1924. Black Males

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	PV 16+	PV[4] 16+ Becker	PV 12	PV 8	(1)-(3)	(2)-(3)	(3)-(4)
a) Actual	47,797.47	n. a.	41,969.65	33,365.43	5,827.82	n. a.	8,604.22
b) Expected	38,024.82	n. a.	29,594.73	23,718.30	8,070.09	n. a.	5,876.43
Ratio $\frac{\text{Actual}}{\text{Expected}}$.					0.72		1.46

Table Ii.
Cohort Age Born 1915-1924. Total Male Population

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	PV 16+	PV[4] 16+ Becker	PV 12	PV 8	(1)-(3)	(2)-(3)	(3)-(4)
a) Actual	81,219.66	81,219.66	62,696.21	47,683.41	18,523.45	18,533.45	15,012.80
b) Expected	79,095.65	82,714.09	57,595.27	38,082.19	21,500.38	25,118.82	19,513.08
Ratio $\frac{\text{Actual}}{\text{Expected}}$,					0.86	0.73	0.77

Notes to Table I:

1. All census data have been converted to constant 1957-1959 dollars.

2. 1940 census data have different educational groupings combining grades 7 and 8, whereas the 1950 and 1960 censuses list grade 8 separately. The 1939 grades 7 and 8 data were considered as grade 8 only.

3. The 1940 census reveals earnings, whereas the two later ones list income. Following the discussion in Becker[1], we have increased all data by 10% to offset the rate of under-reporting of wages and salaries. No adjustment was made to data from 1950 and 1960 censuses because the under-reporting of earnings offset the inclusion of other income, etc.

4. A further adjustment was made to the 16+ data from the 1940 census. Again following Becker[1], the income figures for total male and white male population were increased to offset the effect of the census data not including respondents with "other income" exceeding fifty dollars. The percentage increase varied with age as follows:

Age 25-34: 2.7%
Age 35-44: 6.9%
Age 45-54: 8.6%
Age 55-64: 6.3%

Both the adjusted an unadjusted 16+ data are used in this paper. No adjustment estimates for black males were available.

1. Becker, Gary S., Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, pp. 163 and 168, op. cit.

Chart 1

1939 CROSS-SECTION AND RESULTANT EXPECTED INCOME PATH
FOR TOTAL MALE POPULATION WITH 12 YEARS EDUCATION

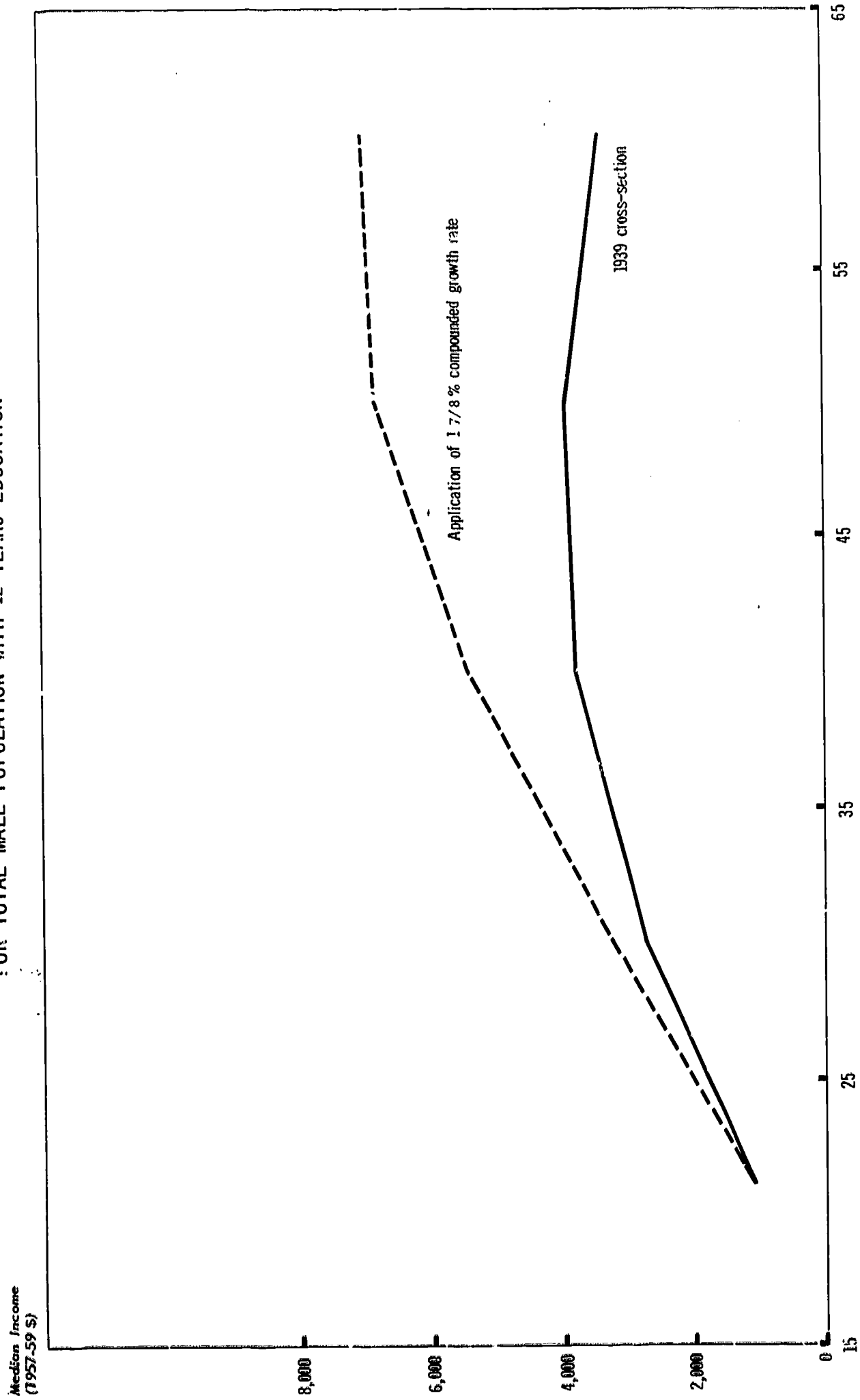


Chart II a

SELECTED ACTUAL INCOME AND EXPECTED INCOME PATHS

COHORT AGE BORN 1915-24 AND 1939 CROSS-SECTION BLOW-UP, TOTAL MALE POPULATION

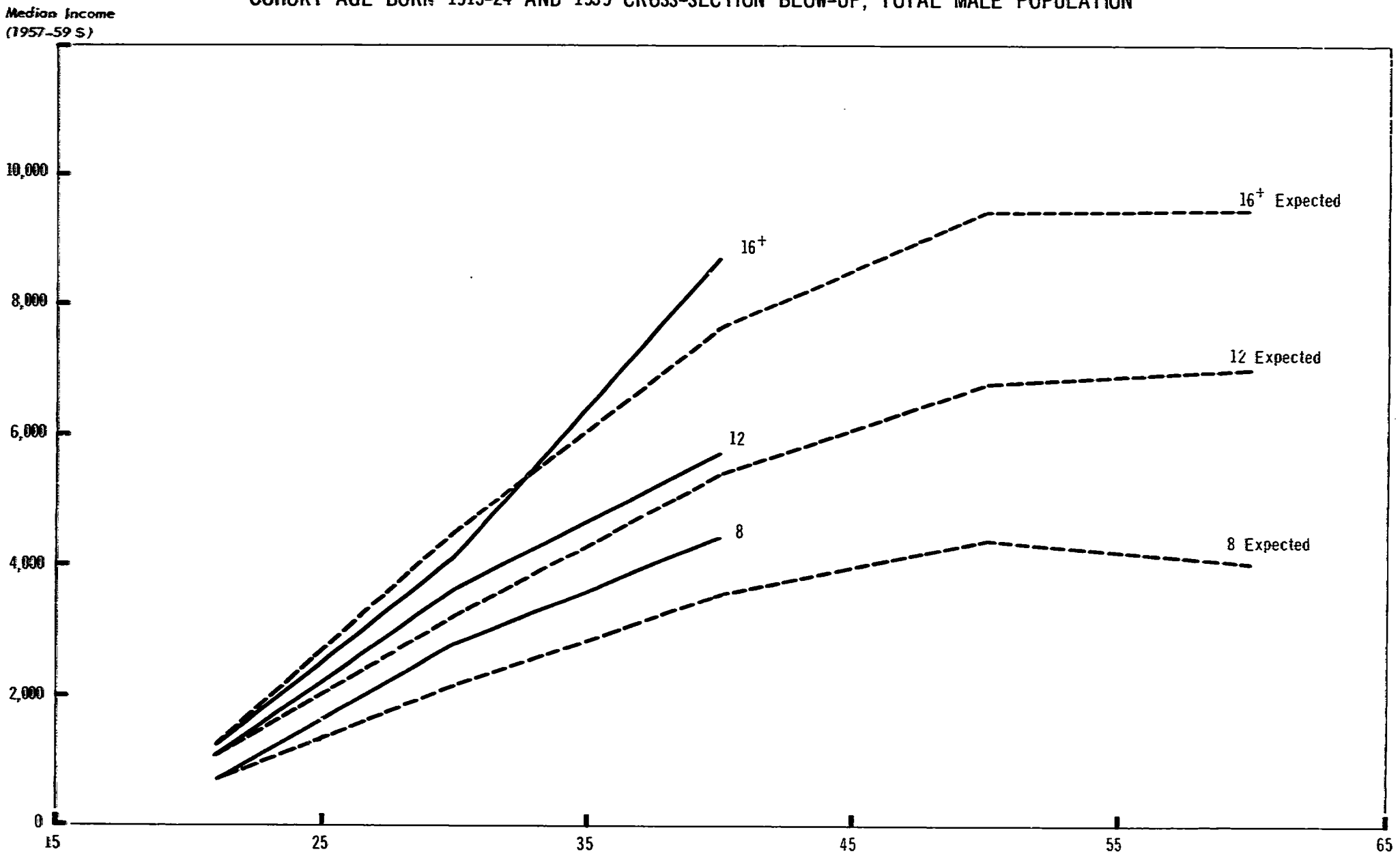


Chart II b

COHORT AGE BORN 1915-24 AND 1939 CROSS-SECTION BLOW-UP, BLACK MALE POPULATION

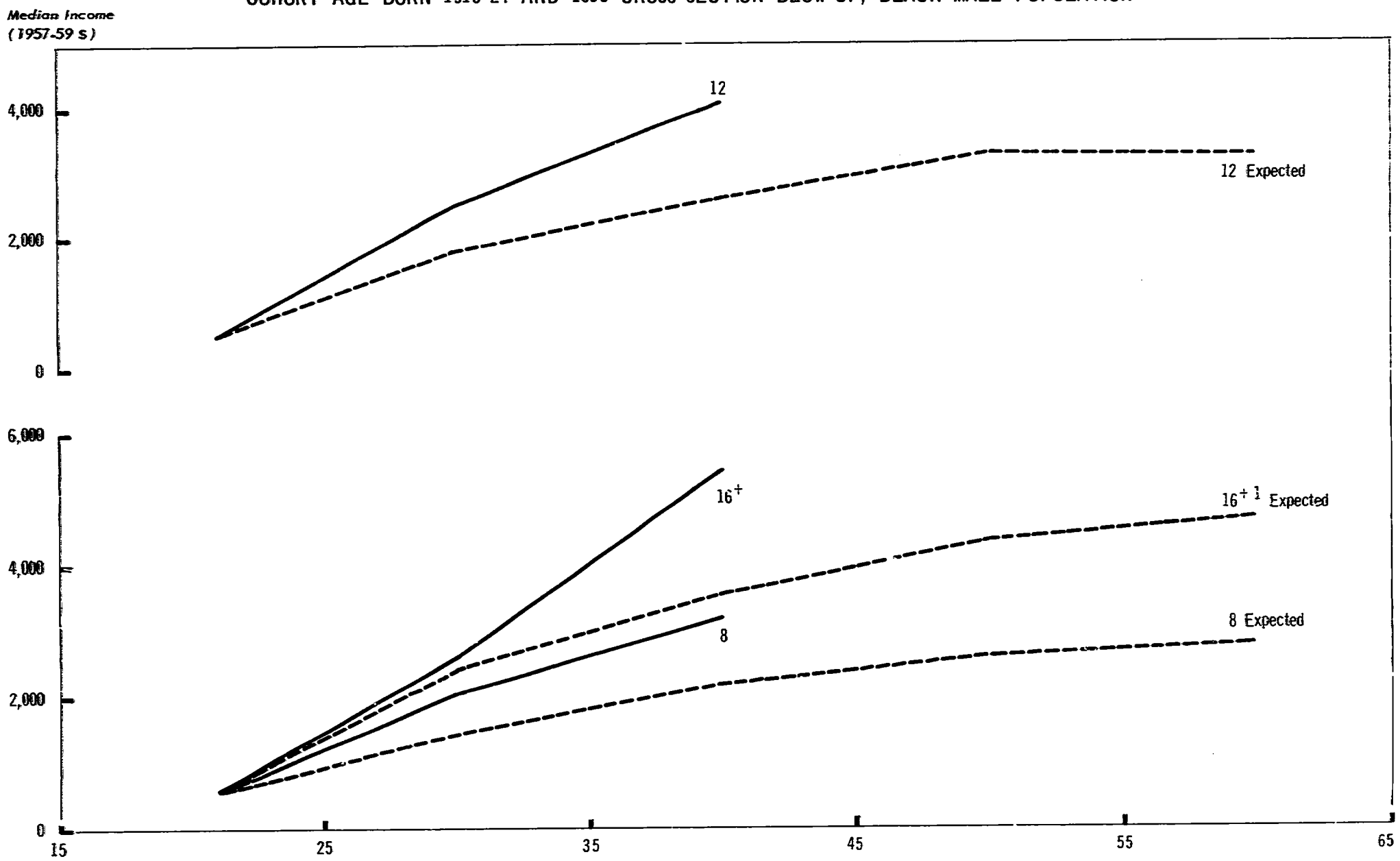


Chart IIc

COHORT AGE BORN 1915-24 AND 1939 CROSS-SECTION BLOW-UP, WHITE MALE POPULATION

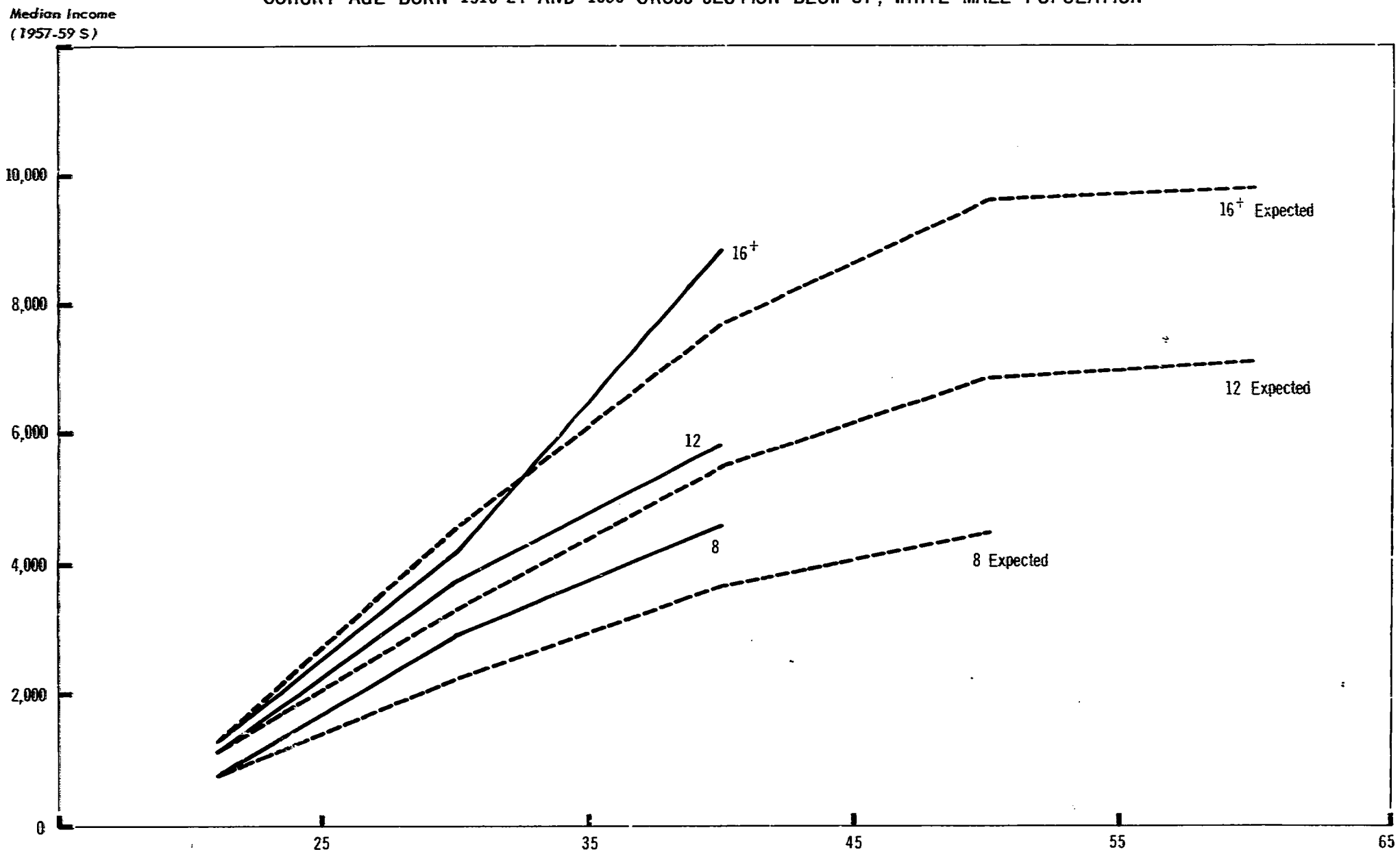


Chart IId

COHORT AGE BORN 1905-14 AND 1939 CROSS-SECTION BLOW-UP, TOTAL MALE POPULATION

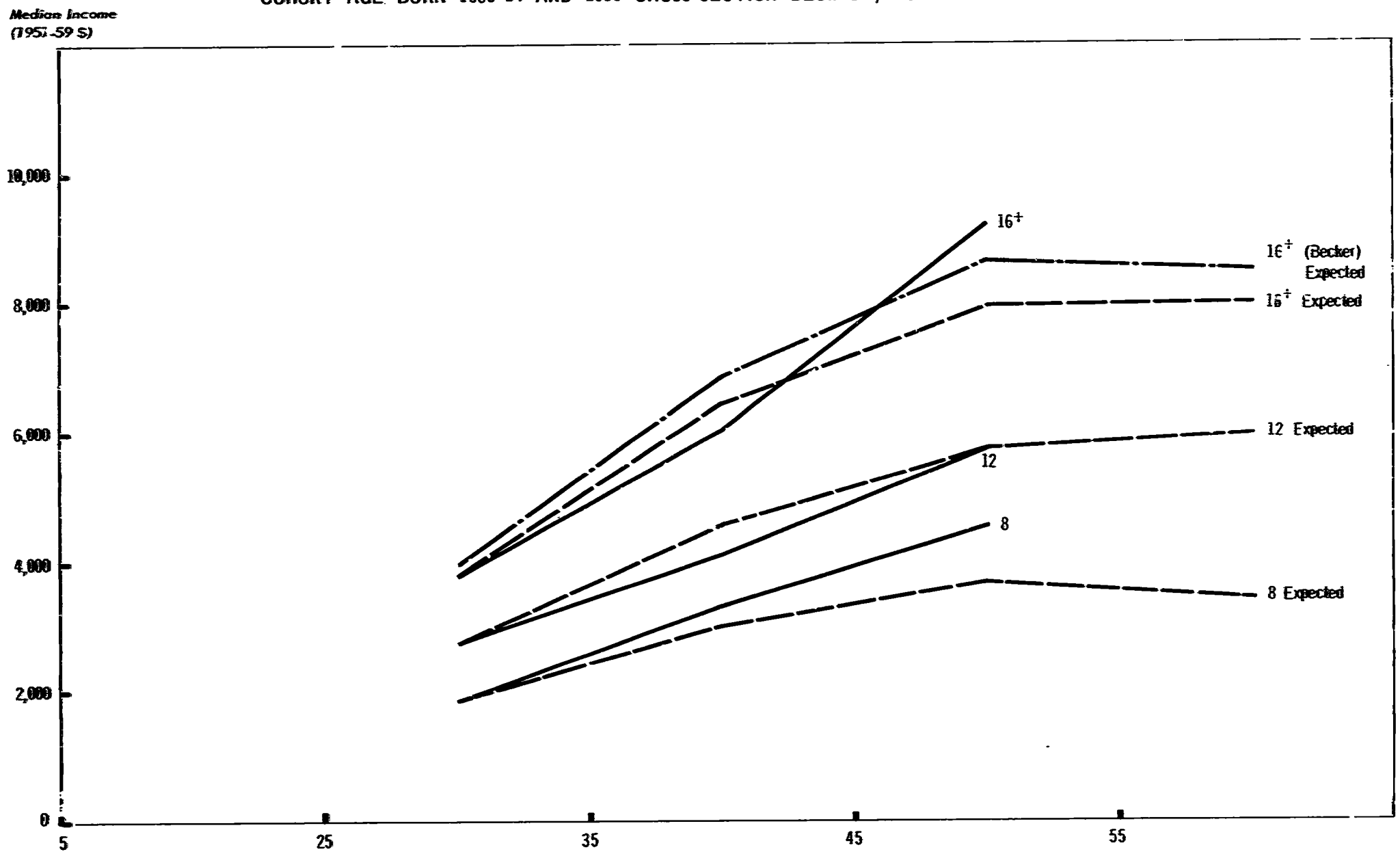


Chart IIe

COHORT AGE BORN 1905-14 AND 1939 CROSS-SECTION BLOW-UP, BLACK MALE POPULATION

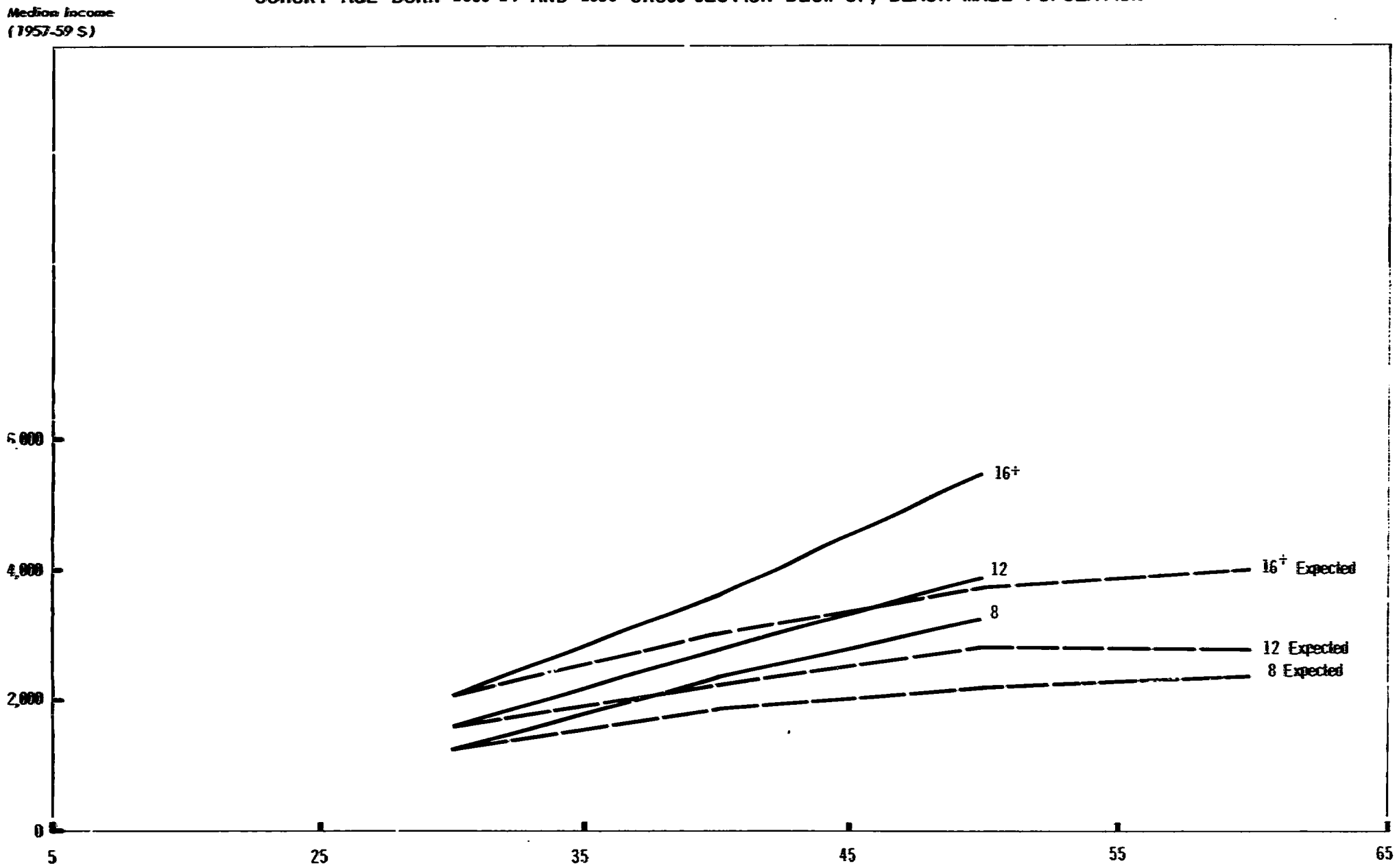


Chart II.f

COHORT AGE BORN 1905-14 AND 1939 CROSS-SECTION BLOW-UP, WHITE MALE POPULATION

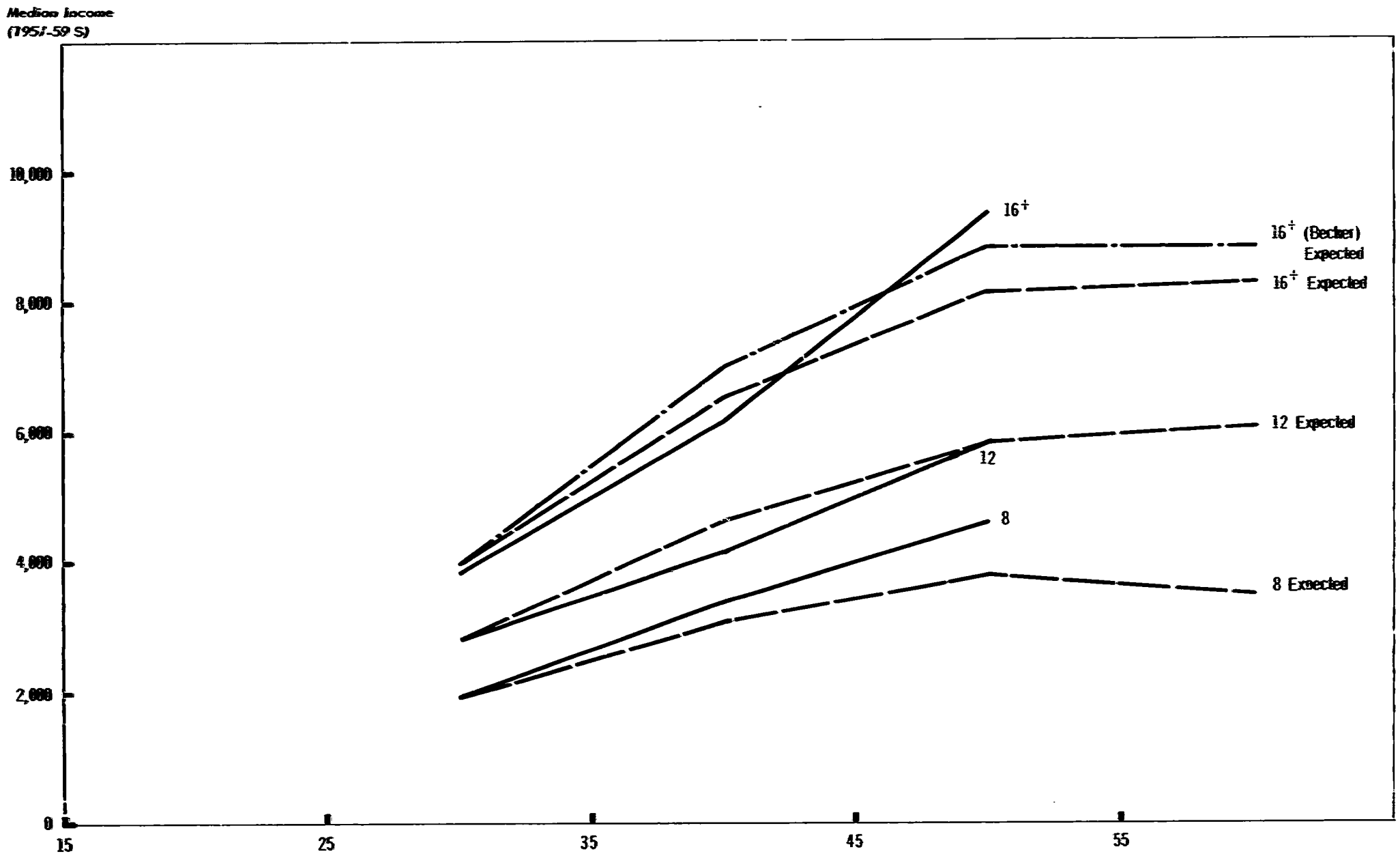


Chart IIg

COHORT AGE BORN 1855-1904 AND 1939 CROSS-SECTION BLOW-UP, TOTAL MALE POPULATION

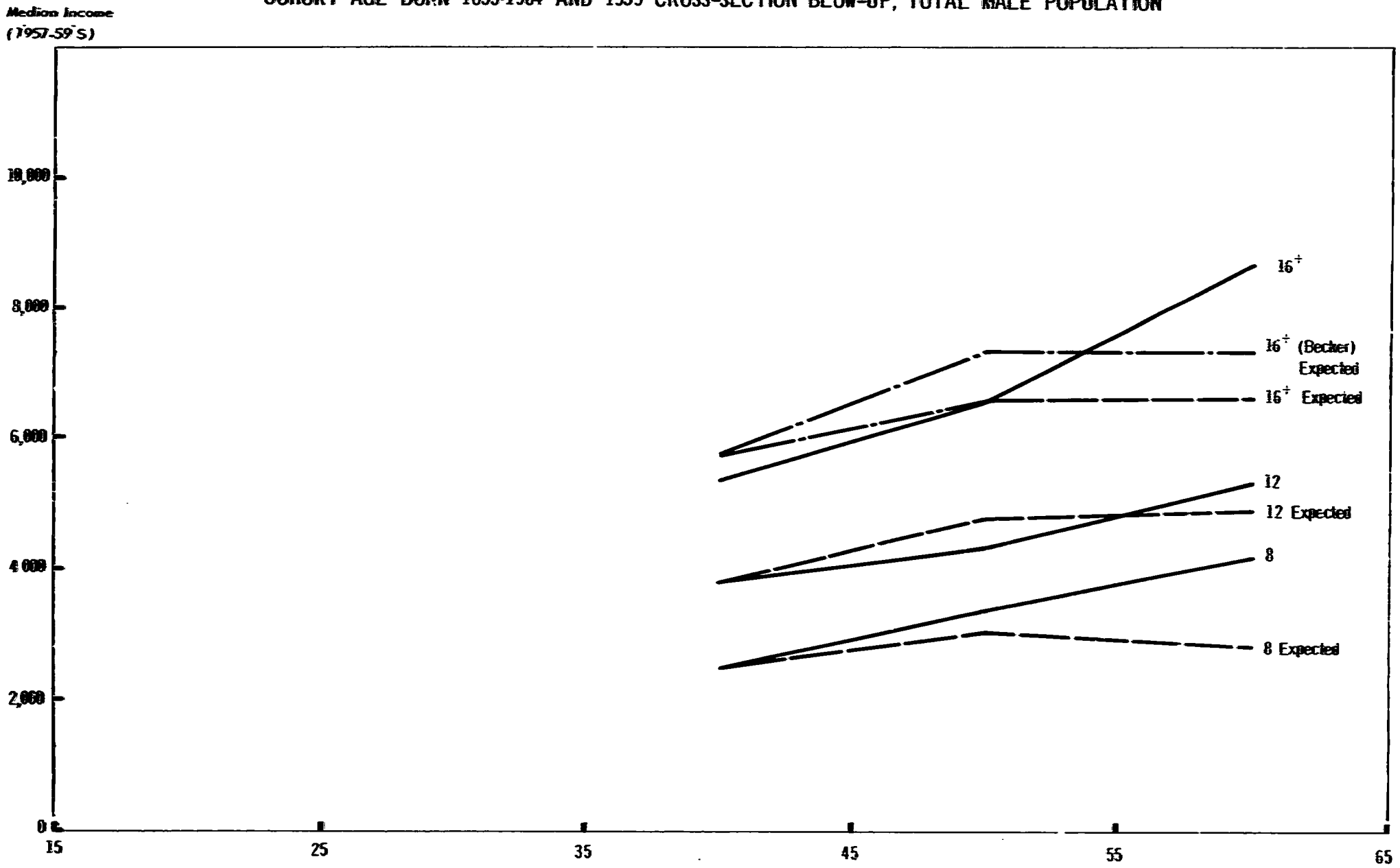


Chart IIh

COHORT AGE BORN 1895-1904 AND 1939 CROSS-SECTION BLOW-UP, BLACK MALE POPULATION

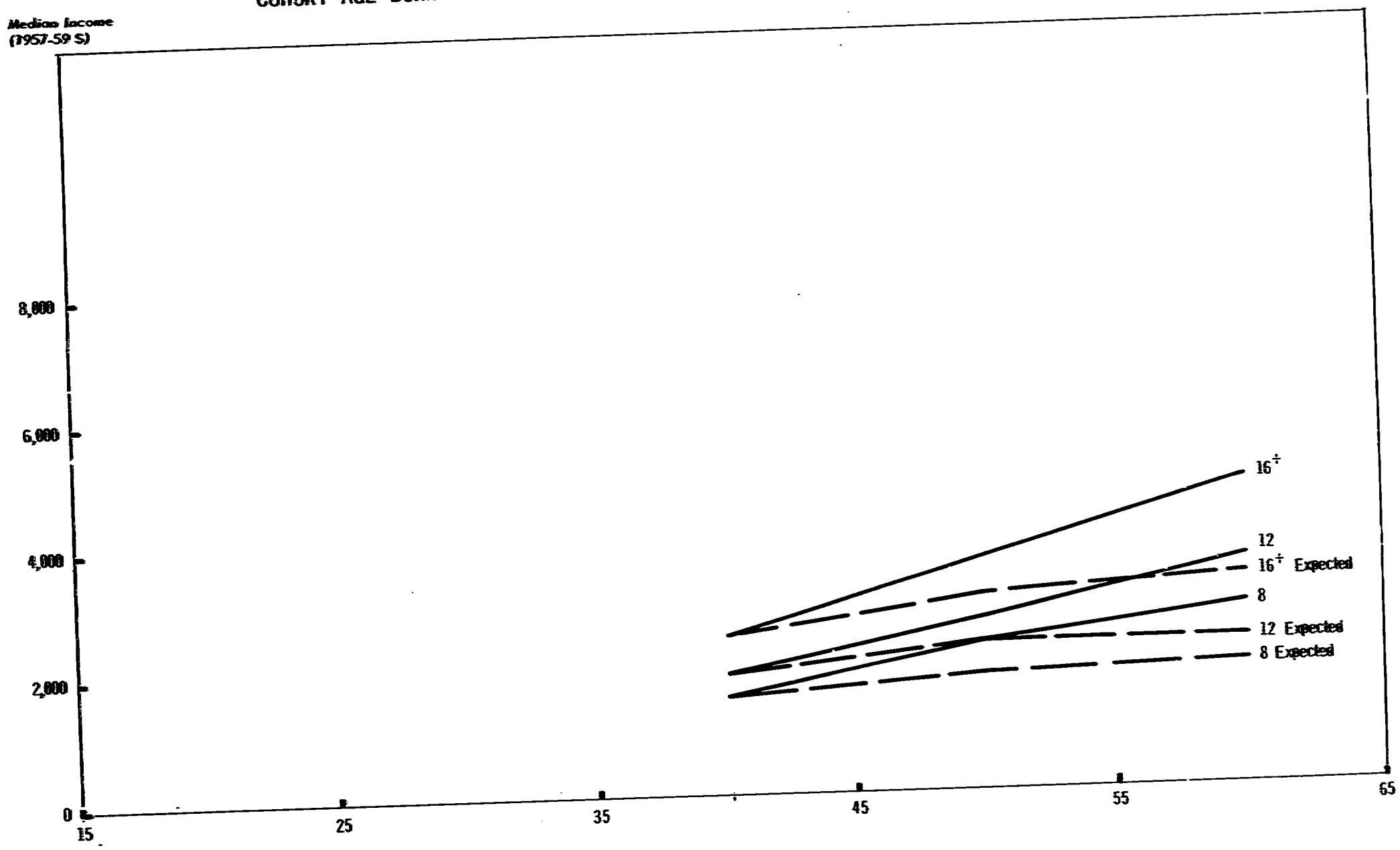
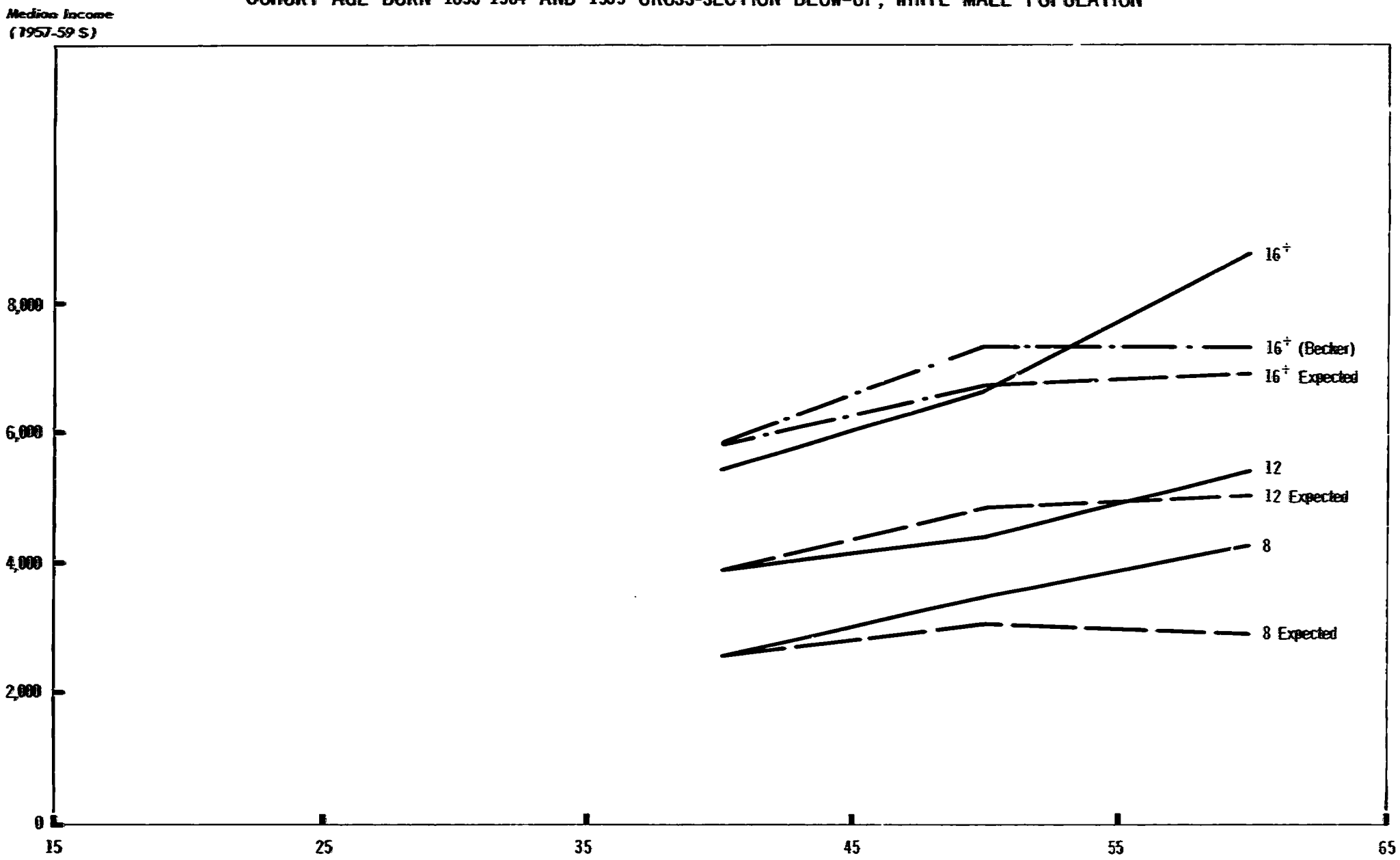


Chart II i

COHORT AGE BORN 1895-1904 AND 1939 CROSS-SECTION BLOW-UP, WHITE MALE POPULATION



A cursory glance at these charts indicates that the actual income paths deviated from those predicted on the basis of the 1939 cross-section and expected growth rate. Moreover, the deviations were different for different levels of education within the same age cohorts, and the pattern of relative deviations for various education levels differed among the different age cohorts. For example, for the total population data, the 1915-1924 age cohort with 12 years of education, the actual income path was somewhat below the expected one, and for the 1895-1904 cohort with 12 years of education, the actual was also below the expected.

We should say a word at this point about the use of the expected rate of growth adjustment. Naturally, the choice of the rate of growth is rather arbitrary; rationales could be provided for somewhat higher or somewhat lower rates of growth of quality-constant labour incomes. However, the arbitrary element in the choice does not affect the conclusions, since the use of the common rate of growth is just a way of providing a common standard between levels of education and across cohorts; the differences in the character of the deviations from the expected path would remain whatever the growth factor used.

Given that there are differences in the character of the deviations of the actual income path from the expected path, it is still somewhat difficult to gather from the charts the significance of such deviations. To do so, let us return to the simple decision situation outlined under Part I. There it was suggested that the present value of additional earnings due to more education might be estimated by taking the present value of the difference between cross-section levels of earnings - increased by a growth factor - and comparing it with the present value of the cost of the additional education. This suggests that the significance of deviations between expected and actual incomes might lie in the difference between the present 1939 value of additional earnings due to education as derived from the expected paths, and the present value as calculated from the actual paths. For example (see Table Ic), for white males in the 1905-1914 cohort, the 1939 present value of income of those with 16+ years of education for the 20-year period up to 1959, as calculated from the expected path, is $72,379. For the same age cohort, the 1939 present value of income for 12 years of education over the same period, as calculated from the expected (i.e. cross-section adjusted for growth) path, is $51,681. Thus the difference in present value of income associated with the extra four years of education, as calculated from the expected paths, is $20,698. Now, the present value for the same cohort over the same period, calculated from actual experience, is $72,950 for 16+ and $48,776 for 12 years of education, giving a difference of $24,174. Thus, the actual difference in present value of income associated with the additional four years was 1.17 times the expected difference in present value, i.e. 17% higher for a period covering about half the working life. This difference seems large enough to be regarded as significant.

When similar calculations of present values for 12 as compared to 8 years of education for the same cohort are made, we find that the actual present value of the difference between 8 and 12 years is only .64 times the expected present value, i.e. 36% less than expected, a significant difference in the opposite direction.

The present values for different cohorts calculated in a way similar to that just described are presented in Tables Ia-i. Some further, but limited, comments on Table I may be in order. First, when one uses the present values for 16+ calculated from paths which use the Becker correction of the 1939 data (i.e. columns (2) and (6) in Tables Ia-i), the deviations of actual from expected present values are smaller for the first two age cohorts (1895-1904 and 1905-1914); in fact, for Table Id and If they are non-existent. However, for the third cohort (1915-1924), the deviations of actual from expected are greater when the Becker correction is used. Thus the correction seems reasonable but does not remove the significance of deviations.

Second, the pattern of deviations for the present value of the additional years between 8 and 12 for white males (and total male population) seems consistent across cohorts, with the actual present value being about .6 to .7 of the expected value. One might conclude from this that 1939 cross-section values for those with 12 years of education were abnormally low. This conclusion would be consistent with the finding for the age cohort 1895-1904 of actual present value for 16+-12 greater than one, but would conflict with the findings for the 16+-12 values for the other two cohorts.

Third, the most striking results are those with respect to the present value of the additional years from 12 to 16+. For the oldest cohort (white and total male), the actual exceeded the expected, for the next cohort it equalled it, and for the youngest cohort it fell considerably short of it.

In most of the literature on the returns to investment in education, evidence is presented in terms of rates of return. It might, therefore, be of interest to give a rough indication of how deviations of actual from expected values of the income stream might affect the rate of return. Appendix A presents the rationale for our calculation of the effects on rates of return, but few typical examples can be given here. If the rate of return based on the expected income stream had been 10%, an actual income stream .63 of the expected (as in column 7, Table Ia) would yield an actual rate of return of 3.9%; an actual/expected of 1.19 (as in column 5, Table Ib) would yield an actual rate of return of 12.9%. The graph in Appendix A can be used to translate other actual/expected figures into rates of return.

One must bear in mind that the estimates of actual incomes are approximative, since they are for ten-year group cohorts and are mapped from only three observations for each actual income path. Just as there were significant differences between 1939 and 1949, there may have been significant differences in opposite directions in the interim years. Since we do not have the information, it is difficult to tell how seriously we may be misled by the fact that we must approximate from only three observations; we do not know, for example, to what extent 1949 was typical. Since we have used medians, we would expect the year-to-year fluctuations to be less than they would have been with means, but the only conclusion - a tentative one - must be that the limited data on hand suggest that there were significant deviations of the actual income experience from what would have been expected on the basis of the simple decision framework outlined in Part I. These deviations differ as between both levels of education and age cohorts.

As was suggested earlier, the question we have been asking is, to what extent have the shifts in the cross-section profiles of age-education-income due to general economic growth been neutral ? Since this concept of neutral shifts is central to the concept of using the cross-section income profiles as proxies for time-series (see Ben-Porath[1]), let us pursue this point a bit further with respect to these 20-year data. We could define neutrality in either of two ways: a) we could look at the age-income path for any given level of education, and call a shift in the cross-section neutral if incomes for each age rose by approximately the same percentage, thus leaving the relative income at different ages about the same; b) we could look at the additional income due to education (e.g. 12 minus 8 years), and call a shift in the cross-section neutral if the differentials at each age rose by the same percentage. Obviously, b) is a somewhat more stringent criterion, but it is the one that seems appropriate to the simple decision framework we have been using.

Charts IIIa - IIIc show the successive cross-sections. Table II gives, for successive cross-sections, the income differential associated with additional education at each age (in percentages). This table represents a test of neutrality of shifts in the sense defined in b). It shows clearly that the shifts were not approximately neutral in this sense.

In Graphs Ia-Ic, we plot the actual income at a given age and education level as a percentage of what it would have been if the previous cross-section had shifted up neutrally in the a) sense at a rate equivalent to the 1-7/8% secular growth rate we have been using. These graphs present then a test of neutrality in the a) sense for individual education levels. If the shifts for any education level had been neutral, the graph would have been approximately a straight line. These graphs show clearly that the test of neutrality in the a) sense is not met for any of the education levels.

In the light of the results obtained by considering the cohort profiles, it is not surprising that the successive cross-sections failed to meet the neutrality of shifts tests. Let us comment on these two different ways of looking at the data. As suggested earlier, these are approximately equivalent ways of posing the same question. If the shifts had been neutral in the a) sense, the deviations of the actual

1. Ben-Porath, Yoram, "Lifetime Income and Economic Growth: Comment", op. cit.

Table II. INCOME DIFFERENTIALS ASSOCIATED WITH ADDITIONAL EDUCATION

(in percentage)

AGE	1939 - 1949		1939 - 1959		1949 - 1959	
	$\Delta(Y_{16+}-Y_{12})$	$\Delta(Y_{12}-Y_8)$	$\Delta(Y_{16+}-Y_{12})$	$\Delta(Y_{12}-Y_8)$	$\Delta(Y_{16+}-Y_{12})$	$\Delta(Y_{12}-Y_8)$
30	-57	- 6	13	16	163	24
40	29	-40	89	- 1	47	63
50	44	-33	129	-17	59	25
60	69	-30	181	-23	66	11

(See note 4 - Table I)

AGE	1939 - 1949	1939 - 1959
	$\Delta(Y_{16+}\text{Becker}-Y_{12})$	$\Delta(Y_{16+}\text{Becker}-Y_{12})$
30	-61	2
40	4	53
50	10	75
60	36	126

from the expected cohort path would have been systematic for a given level of education across cohorts. If they had been neutral in the b) sense, the deviations in actual and expected paths and the differences in present value would have been systematic across both cohorts and levels of education. However, we feel that it is preferable to analyse the data in terms of the cohort profiles. Looking at successive cross-sections as a whole can be somewhat misleading. We tend to forget, for example, that a present value made up from a cross-section is a weighted (by the discount rate) sum of yearly incomes. The present value actually experienced by a cohort is going to be drawn from different components of successive cross-sections. The fact that these are different weighted sums of components means that it is not easy to deduce from the cross-sections as a whole the actual experience of particular cohorts. Once the shifts in successive cross-sections deviate from neutrality in either sense, the relationship between the cross-section profiles and the actual cohort experience becomes rather complex. One might conjecture that the propensity to concentrate on cross-section profiles as a whole in the past has tended to obscure some of the issues we are exploring here.

We can summarize this section by concluding that the twenty-year data indicate that the actual experience of groups with different levels of education deviated significantly from what would have been predicted at the beginning of the period on the basis of the cross-section data available at that time. Had educational investment decisions been made according to the simple framework outlined under Part I, the actual returns on investment realized over the twenty years would have been substantially different

Chart III

SUCCESSIVE CROSS-SECTIONS FOR TOTAL MALE POPULATION

Chart IIIa

1939 AND 1949 CROSS-SECTIONS

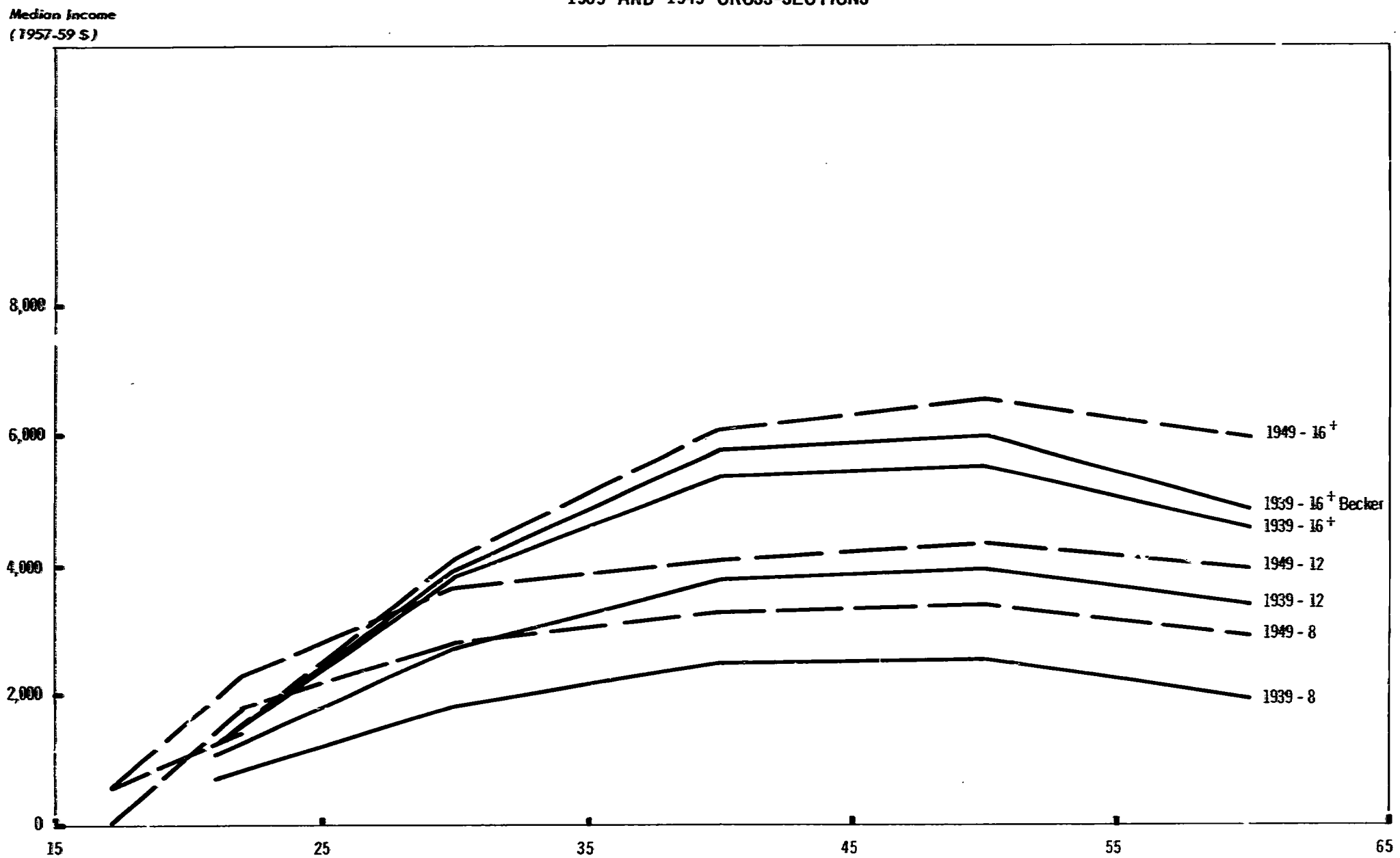


Chart III b

1939 AND 1959 CROSS-SECTIONS

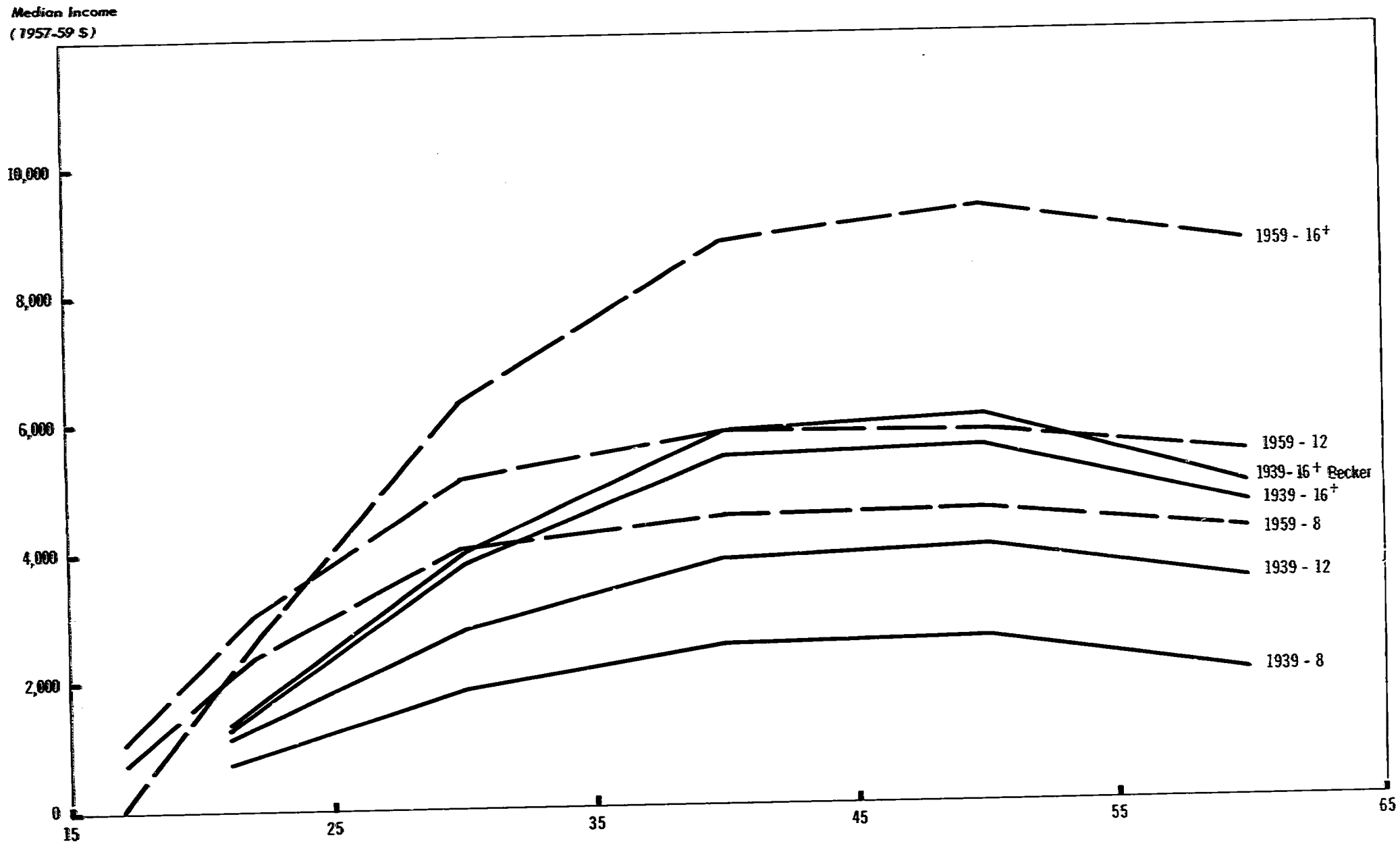
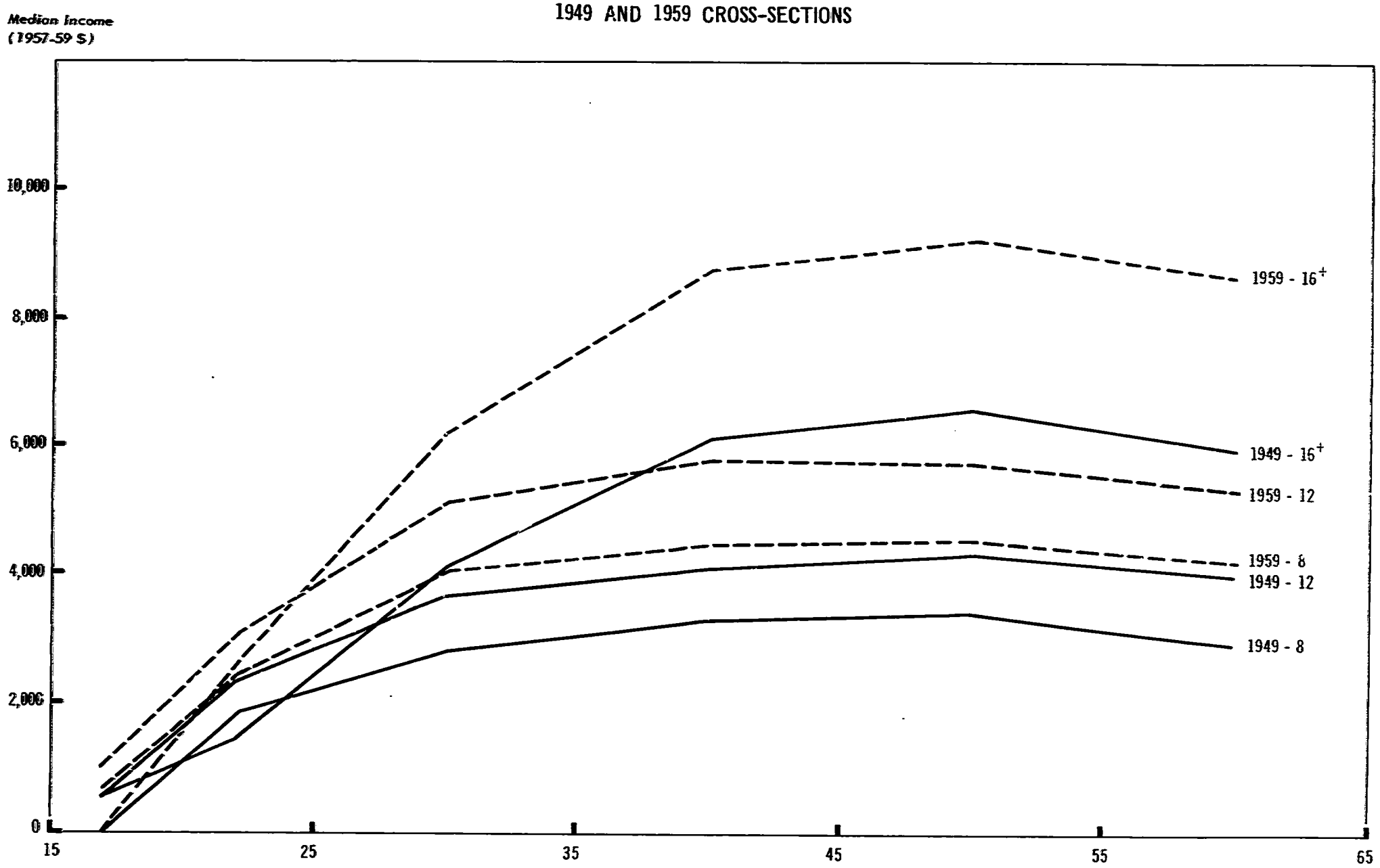


Chart IIIc

1949 AND 1959 CROSS-SECTIONS

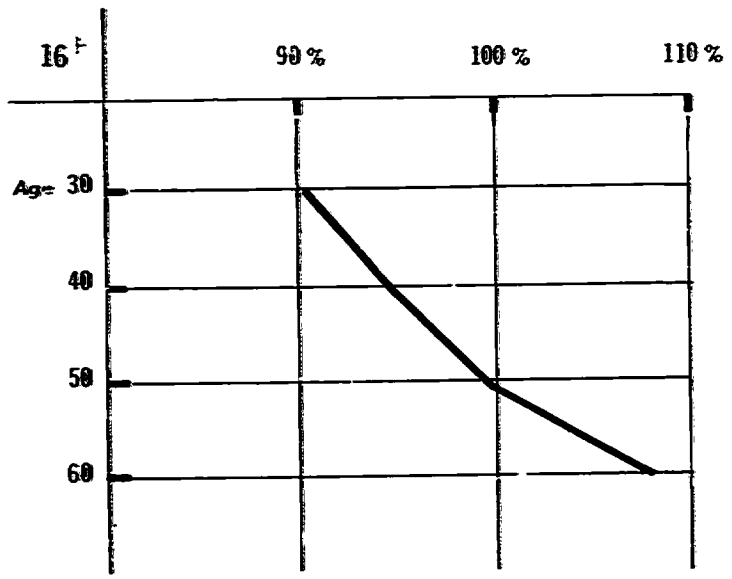


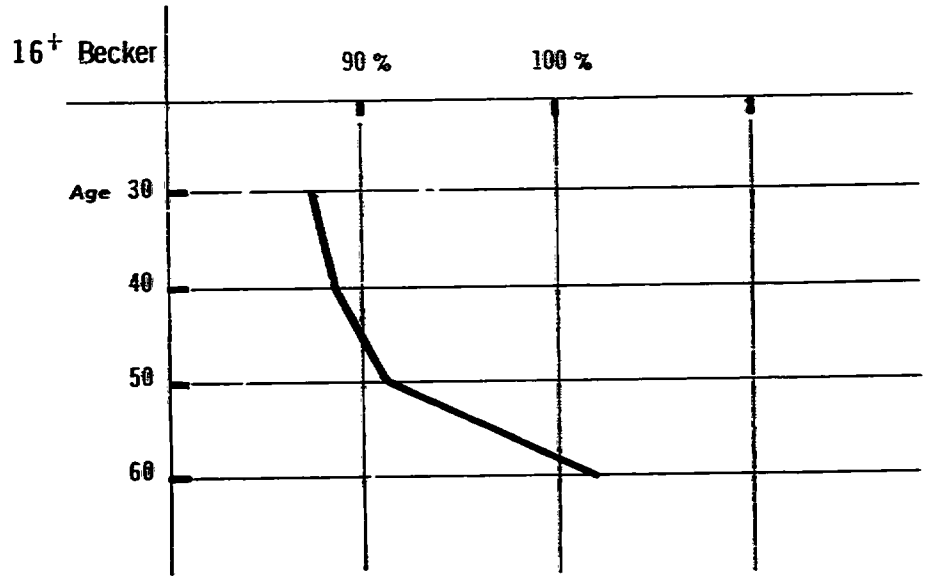
Graph I

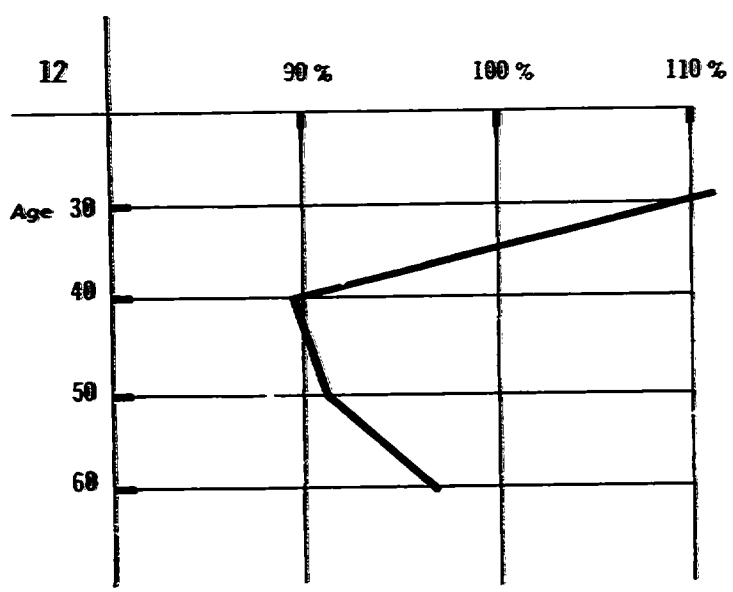
ACTUAL INCOME AS A PERCENTAGE OF CROSS-SECTION NEUTRAL GROWTH

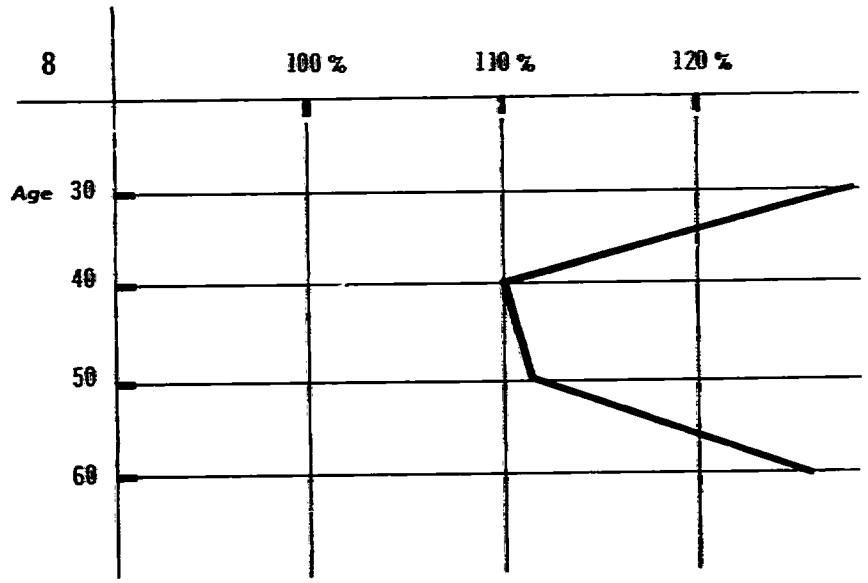
Graph Ia

1949 ACTUAL INCOME/1939 NEUTRAL GROWTH



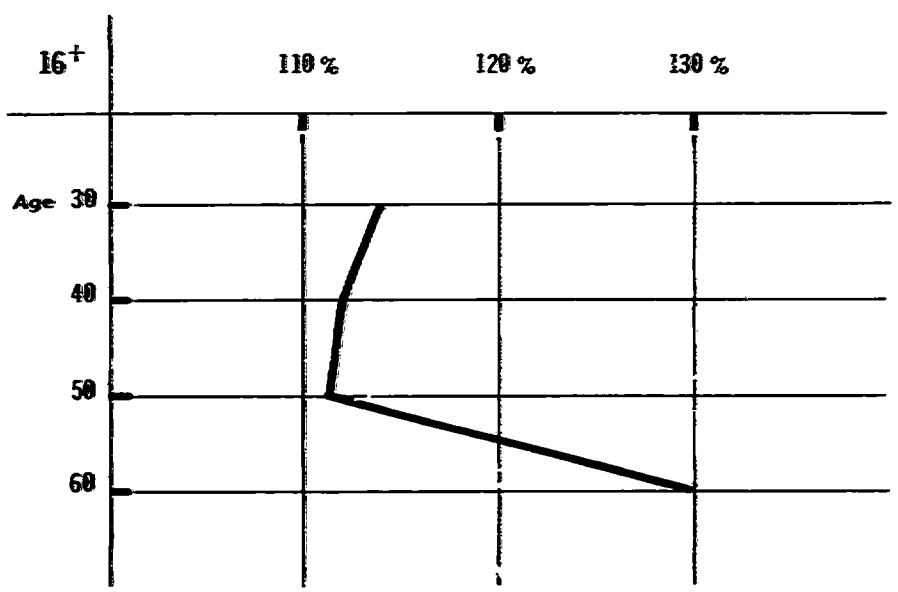


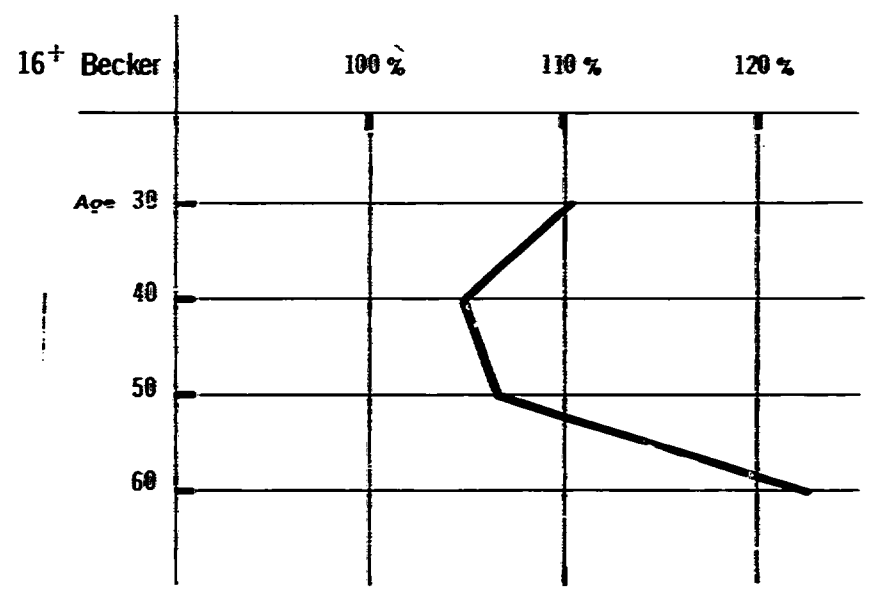


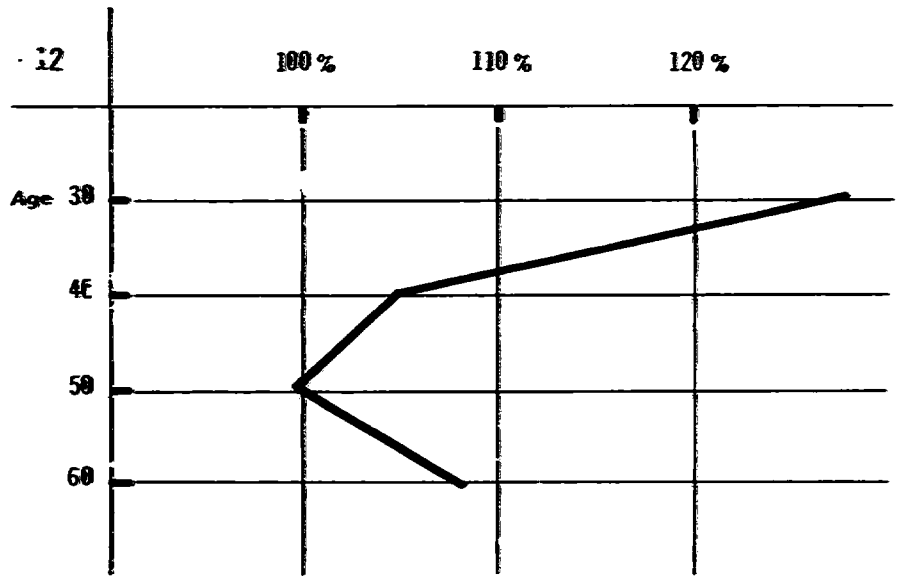


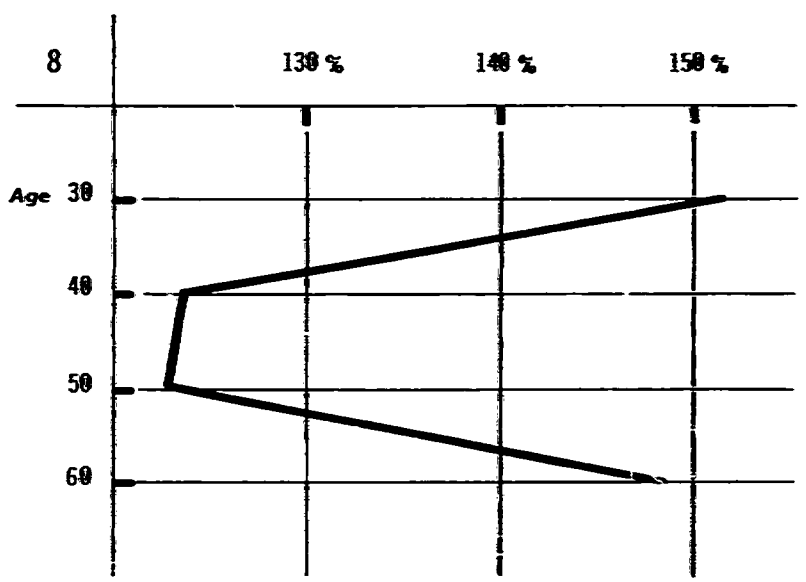
Graph 1b

1959 ACTUAL INCOME/1939 NEUTRAL GROWTH



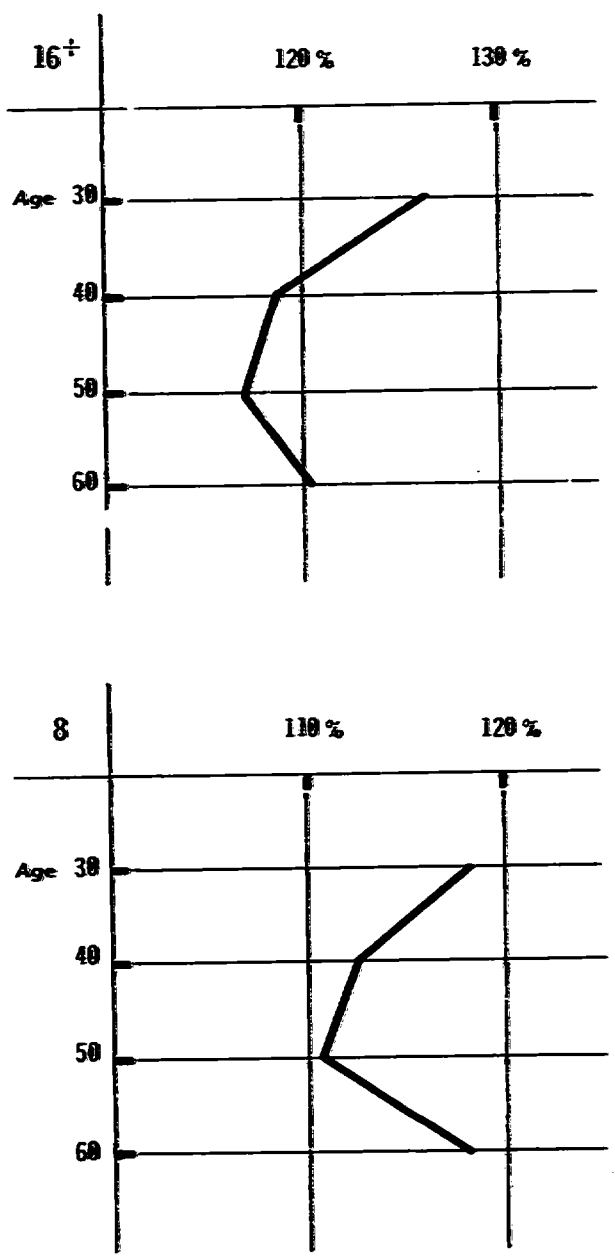


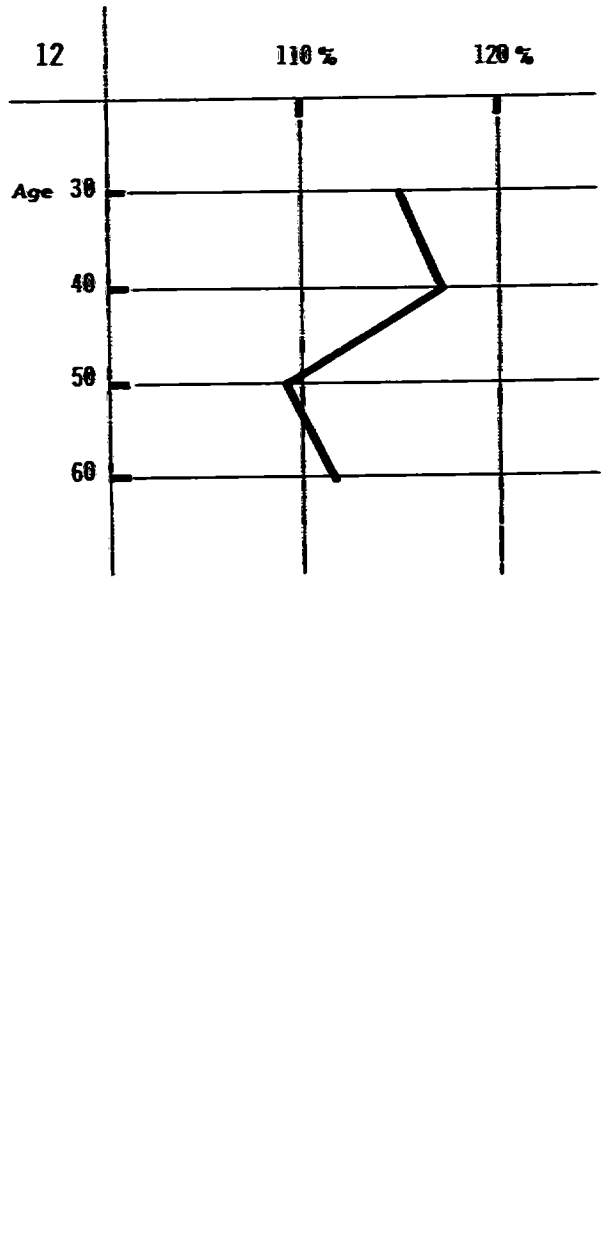




Graph Ic

1959 ACTUAL INCOME/1949 NEUTRAL GROWTH





from those expected. It should be stressed that this conclusion is based upon very scant data, and thus must be considered as a very tentative one.

Unfortunately, this finding does not advance us very much in the application of human capital theory to educational investment decisions. It suggests that we should reject the rather simple decision framework outlined earlier, and seek to penetrate the greater complexity of inter-relationships alluded to at the beginning of this paper.

III

THE 1956 TO 1966 EXPERIENCE

There is another body of data available which permits a somewhat more detailed, though still limited, exploration of the problems with which we are concerned here. A Current Population Report of the United States Bureau of the Census[1] gives age-education-income data for men in the United States for selected years between 1956 and 1966. Since these data are provided for six years of this interval and for individual age cohorts, more analysis is possible even though the period is shorter.

We concluded earlier that, since the present values expected on the basis of the 1939 cross-section differed significantly from the actual present values as experienced by cohorts (cross-sections did not shift in a neutral fashion between 1939 and 1949 and 1959), an attempt should be made to analyse the possible effect of supply changes, demand shifts and cyclical effects. We have not tried to construct an elaborate model which would permit us unequivocally to separate out these various factors. Our approach here is much more eclectic, and, in any case, we shall take several different analytical tacks in trying to obtain further insight into the relationships between age, education and income.

As a first step in the analysis of these data, we have repeated the earlier calculation of expected and actual present values for various age cohorts and levels of education. We have used, for this purpose, the cross-section observations on mean income for each age and education level for 1956 and multiplied these values by the appropriate compounded value of the growth rate (using the same value as before) in order to construct an expected income path in a fashion similar to that illustrated in Diagram I.

The results of these calculations are reported in Tables IIIa-d. It should be noted that, with these data, it has been possible to trace expected and actual income paths for single-year age groups (instead of the ten-year groups used before). Four different age cohorts were selected for analysis: those who were respectively 50, 40, 35, and 30 in 1956.

Once again we find sizeable deviations of the actual from the expected present values; the ratios of actual to expected values range for the various cohorts from .86 to 1.02 for the 16+ - 12 years, and from .96 to 1.20 for the 12 - 8 years. Again, using the method outlined in Appendix A, we see that, if the expected rate of return were 10%, then, for an actual/expected equal to .86, the actual rate of return would be 6.4%; and that for an actual/expected of 1.20, it would be 14.5%.

It is again noteworthy that the experience of different age cohorts over the same ten years, 1956-1966, is quite different; for each cohort, the extent to which the actual experience deviated from what would have been predicted from the 1956 cross-section was somewhat different. Only two general patterns emerge: first, with the exception of the oldest cohort (1906), the ratio of actual to expected was on the

1. United States Bureau of the Census, Department of Commerce, "Annual Mean, Lifetime Income and Educational Attainment of Men in the United States for Selected Years 1956 to 1966", Current Population Reports, Consumer Income, Series p-60, No. 56, August 1968.

Table III. PRESENT VALUES - EXPECTED AND ACTUAL[1]
(IN CONSTANT 1957-1959 DOLLARS)
CALCULATED AT A DISCOUNT RATE OF 6% FOR THE TEN YEARS 1956-1966

Table IIIa. Total Males, Cohort Age Born 1906

	(1)	(2)	(3)	(4)	(5)
	PV 16+	PV 12	PV 8	(1) - (2) 16 - 12	(2) - (3) 12 - 8
a) Expected	32,108.02	18,792.08	13,246.05	13,316.12	5,546.03
b) Actual	31,886.34	19,012.96	12,846.02	12,873.38	6,166.94
Ratio $\frac{\text{Actual}}{\text{Expected}}$				0.968	1.11

Table IIIb. Total Males, Cohort Age Born 1916

	(1)	(2)	(3)	(4)	(5)
	PV 16+	PV 12	PV 8	(1) - (2) 16 - 12	(2) - (3) 12 - 8
a) Expected	36,960.32	21,200.79	15,578.66	15,759.53	5,622.13
b) Actual	34,293.40	20,687.63	14,927.06	13,605.77	5,660.57
Ratio $\frac{\text{Actual}}{\text{Expected}}$				0.862	0.99

Table IIIc. Total Males, Cohort Age Born 1921

	(1)	(2)	(3)	(4)	(5)
	PV 16+	PV 12	PV 8	(1) - (2) 16 - 12	(2) - (3) 12 - 8
a) Expected	46,132.17	28,691.51	21,357.90	17,440.66	7,333.61
b) Actual	43,419.64	27,367.44	20,297.95	16,052.20	7,069.49
Ratio $\frac{\text{Actual}}{\text{Expected}}$				0.919	0.96

Table IIId. Total Males, Cohort Age Born 1926

	(1)	(2)	(3)	(4)	(5)
	PV 16+	PV 12	PV 8	(1) - (2) 16 - 12	(2) - (3) 12 - 8
a) Expected	53,494.80	37,112.92	28,020.67	16,381.88	9,192.25
b) Actual	52,495.23	35,320.51	26,475.10	17,174.72	10,845.41
Ratio $\frac{\text{Actual}}{\text{Expected}}$				1.022	1.2

1. Data for these calculations were drawn from "Annual Mean, Lifetime Income and Educational Attainment of Men in the United States for Selected Years 1956 to 1966" Current Population Reports, Consumer Income, Series P-60 No. 56, 14th August, 1968, Bureau of the Census, United States Department of Commerce.

same side of one for both levels of education (16+ - 12 and 12 - 8); second, in every cohort, the ratio of actual to expected was higher for the 12 - 8 level of education than for the 16+ - 12 level. We shall not comment further on these present value results, because they are primarily intended to establish the fact that sizeable deviations of actual from expected present values were observed for this period; analysis of possible causes of such deviations is more readily carried out with different configurations of the data.

The question of the relationship between cross-section income profiles and actual cohort income experience has been studied by another author in a somewhat different context. H. Miller[1] examined the influence of economic growth on the incomes of various education cohorts over the period 1950 to 1960. He compared the actual income change from 1950 to 1960 for each cohort, with the appropriate income difference in the 1950 cross-section profile, and concluded: "When the age component of the total increase is taken into account, it appears that economic growth accounted for a 5.1% annual increase in income between ages 30 and 40; a 2.0% increase between ages 40 and 50; and 1.8% increase between ages 50 and 60. Since the latter element is not taken into account in traditional measures of estimating lifetime income, it appears that its inclusion would add to the expected income gains of younger men and would therefore have an important bearing on the estimates of expected lifetime income."

Ben-Porath[2] suggested a correction to Miller's procedures which would eliminate the interaction between movement along the cross-section profile and a shift in that profile over time. After eliminating this interaction, Ben-Porath concludes: "The pure growth effect (eliminating the interaction) is somewhat higher in the age group 35-44 than in older age groups, but the decline where it exists is much smaller than reported by Miller, and in some cases there is actually an increase with age where Miller shows a decline. It is also interesting to note that the group least affected by growth is the group 45-54 and not the oldest group 55-64."

Now these studies represent at least an attempt to get at some of the factors which might cause cohort experience to deviate from cross-section profiles. The authors were suggesting that "economic growth" was distorting the cross-section profile. As was suggested earlier in this essay, a variety of factors could cause such non-neutral shifts in the income profile. We shall return later to a discussion of such factors but, first, let us compare the Miller and Ben-Porath results with those obtained from the data we have been using[3].

For this purpose, we examine the data reported in Tables IVa-d. For the same age cohorts selected for analysis in Table III and for the same period 1956-1966, we report in column 2 of the tables, for each of the years of observation, the ratio of the actual income of the education-age cohort to the expected income for that age-education class (constructed as described above). If we take the last entry in that column for each cohort education level, we have an indication of the experience of each group at the end of the ten-year period. This gives us a basis for comparison with the Miller and Ben-Porath data, which also covered a ten-year period (1950-1960). We find that for the 16+ level, the cohort with the highest actual income relative to expected income is the oldest one, 1906 (who were 50 to 60 years of age during the period). Next was the yougest cohort, 1926, followed by the 1921 and, lastly, by the 1916 cohort, which had an actual income some 5% below its expected income. At first glance, this would seem roughly in line with Ben-Porath's conclusions, the youngest and the oldest cohorts doing better than those in between. But these conclusions do not follow as one moves down to the lower levels of education: one of the middle age cohorts, 1916, does best in the actual/expected for level 12 and level 8, while the oldest cohort does worse at level 12 and the youngest cohort does worse at level 8.

1. Miller, Herman P., "Lifetime Income and Economic Growth", op. cit.
2. Ben-Porath, Yoram, "Lifetime Income and Economic Growth: Comment", op. cit.
3. Our method of constructing an expected income path and then checking for deviations from it is roughly equivalent to Ben-Porath's method of allowing for interactions.

Table IV. RELATION OF ACTUAL MEAN INCOME
TO EXPECTED MEAN INCOME
FOR SELECTED AGE COHORTS FOR YEARS 1956-1966

(Same Data Base as Table III)

Table IVa.
Cohort Age Born 1906

AGE	RATIO $\frac{\text{ACTUAL}}{\text{EXPECTED}}$
16+ YEARS OF EDUCATION	
50	1.0000
52	.9901
55	.9448
57	1.0379
58	.9702
60	1.0890
12 YEARS OF EDUCATION	
50	1.0000
52	1.1786
55	.9475
57	.9626
58	.9536
60	.9614
8 YEARS OF EDUCATION	
50	1.0000
52	.9142
55	1.0145
57	.9558
58	.9538
60	.9919

Table IVb.
Cohort Age Born 1916

AGE	RATIO $\frac{\text{ACTUAL}}{\text{EXPECTED}}$
16+ YEARS OF EDUCATION	
40	1.0000
42	.9045
45	.9285
47	.9017
48	.8991
50	.9558
12 YEARS OF EDUCATION	
40	1.0000
42	.9599
45	.9486
47	.9858
48	.9733
50	1.0132
8 YEARS OF EDUCATION	
40	1.0000
42	.9122
45	.9709
47	.9687
48	.9246
50	.9951

Table IV. RELATION OF ACTUAL MEAN INCOME
TO EXPECTED MEAN INCOME
FOR SELECTED AGE COHORTS FOR YEARS 1956-1966

(Same Data Base as Table III)

Table IVc.
Cohort Age Born 1921

AGE	RATIO $\frac{\text{ACTUAL}}{\text{EXPECTED}}$
16+ YEARS OF EDUCATION	
35	1.0000
37	.9136
40	.9471
42	.9192
43	.9225
45	.9651
12 YEARS OF EDUCATION	
35	1.0000
37	.8866
40	.9374
42	.9827
43	.9599
45	1.0044
8 YEARS OF EDUCATION	
35	1.0000
37	.9084
40	.9416
42	.9582
43	.9351
45	.9908

Table IVd.
Cohort Age Born 1926

AGE	RATIO $\frac{\text{ACTUAL}}{\text{EXPECTED}}$
16+ YEARS OF EDUCATION	
30	1.0000
32	.9981
35	.9899
37	.9391
38	.9565
40	1.0014
12 YEARS OF EDUCATION	
30	1.0000
32	.8986
35	.9349
37	.9747
38	.9552
40	.9863
8 YEARS OF EDUCATION	
30	1.0000
32	.9034
35	.9244
37	.9530
38	.9459
40	.9816

One might conclude that the differences between our results and those of Miller and Ben-Porath are due to differences in the source of the data and the lack of complete overlap in the periods (1956-66 vs. 1950-60). But the implications of these findings go further than that. Table V gives the relative ranks of age cohorts on the basis of the ratio reported in column 2 of Tables IVa-d. The first column shows the ranks for the last year of observation, 1966, the results just described in the previous paragraphs. The next column shows the ranks for the middle year of the period, 1961.

Table V. RANKS OF COHORTS ACCORDING TO ACTUAL INCOME/EXPECTED INCOME

YEAR 1966	YEAR 1961
LEVEL 16+	LEVEL 16+
1. Cohort 1906	1. Cohort 1926
2. Cohort 1926	2. Cohort 1921
3. Cohort 1921	3. Cohort 1906
4. Cohort 1916	4. Cohort 1916
LEVEL 12	LEVEL 12
1. Cohort 1916	1. Cohort 1916
2. Cohort 1921	2. Cohort 1906
3. Cohort 1926	3. Cohort 1921
4. Cohort 1906	4. Cohort 1926
LEVEL 8	LEVEL 8
1. Cohort 1916	1. Cohort 1906
2. Cohort 1906	2. Cohort 1916
3. Cohort 1921	3. Cohort 1921
4. Cohort 1926	4. Cohort 1926

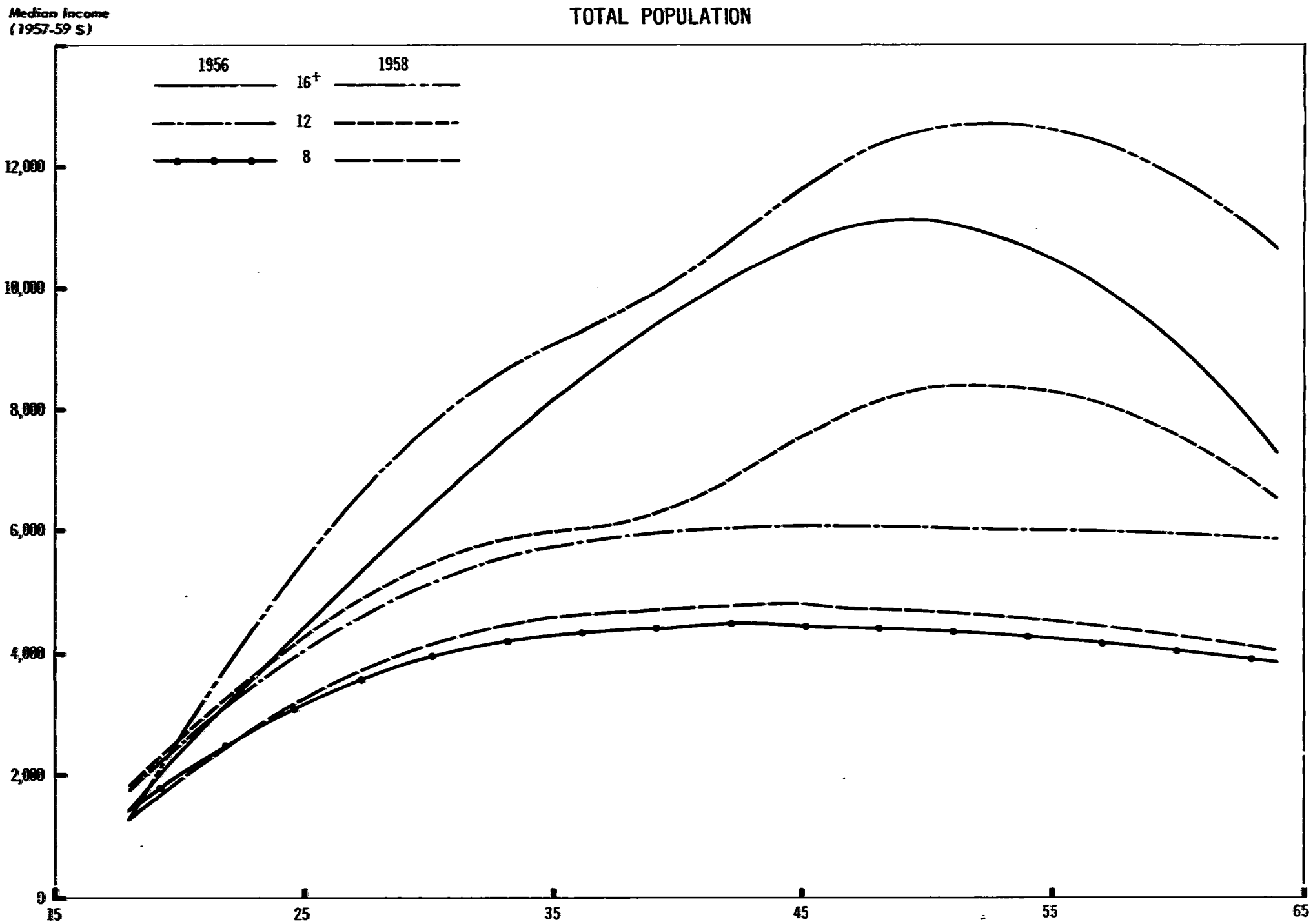
We can see from this table that rankings not only differ according to the level of education, but also change considerably when a different terminal year is chosen as the point of comparison. This suggests that lumping shifts in the profile over a period into a broad term such as "the effect of economic growth" is misleading. The fact that the conclusions one might draw about such "effects" are likely to be very sensitive to the choice of the terminal year, suggests that the effects observed are due to something more than a broad and steady force of the type one imagines when the term "effects of economic growth" is used. The figures reported in column 2 of Tables IVa-d give the actual/expected income ratio for each period, so that it is possible to see the extent to which the relative performance of each cohort education level group changed from year to year.

The judgment of the relative effects of the shifts over time on various cohorts is affected not only by the choice of the terminal year, but even more basically by the choice of the initial year. The latter is particularly important because it is that year's cross-section profile which provides the basis for the construction of the expected income path for all the cohorts. The importance of the choice of the base year is easily seen by looking at the cross-section profiles shown in Chart IV, where the 1958 cross-section is shown in relation to the 1956 cross-section. Table VI shows the actual/expected income ratios for 1966 which result when the expected income is constructed on the basis of the 1958 cross-section.

Chart IV

CROSS-SECTION - 1956

TOTAL POPULATION



A comparison of the results in Table VI with the corresponding figures in Table IV (or comparing rankings from Table VI with those reported in column (1) of Table V) shows that the choice of a different base year alters considerably the conclusions. For example, if we compare the 1916 and 1926 cohorts at level 16+ using the 1956 cross-section base (Table IV), we see that, at the end of the period, the 1916 cohort has an actual/expected 5% worse than the 1926 cohort, whereas, if we use the 1958 cross-section base(Table VI),the 1916 cohort has an actual/expected 5% better. Similar differences in results appear across education levels with a given cohort as well as across cohorts within a given education level.

Table VI. $\frac{\text{ACTUAL}}{\text{EXPECTED}}$ INCOME IN 1966 - USING 1958 CROSS-SECTION TO CONSTRUCT EXPECTED INCOME

(Same data base as Table III)

COHORT	1906	1916	1921	1926
Level 16+	1.0999	1.0568	1.0564	1.0034
Level 12	0.8157	1.0555	1.1328	1.0976
Level 8	1.0850	1.0909	1.0907	1.0866

One simple explanation of the differences between the results when using 1956 or 1958 as a base immediately springs to mind: 1958 was a year of very high unemployment. In fact, there were rather considerable fluctuations in the unemployment rate throughout the 1956-1966 period, and these are reported in Table VII. The differential effects of fluctuations in the level of economic activity on various age-education levels could have a lot to do with the variations in the experience of the various cohorts.

Table VII. UNEMPLOYMENT RATES 1956-1966

YEAR	%
1956	4.1
1957	4.3
1958	6.8
1959	5.5
1960	5.5
1961	6.7
1962	5.5
1963	5.7
1964	5.2
1965	4.5
1966	3.8

All this amounts to saying that deviations of actual from expected income of various cohorts cannot be explained by the effects of some broad, simple trends. To gain a better understanding of the forces which are affecting the lifetime incomes of various cohort education groups, it would seem necessary to go beyond simple comparisons of trends and to formulate more complex analytical models. The next part describes some attempts in this direction. Let us say at once that we do not regard these attempts as much of an advance - the formulation of the model is rather haphazard, and the limitations of the data seriously restrict the validity of any inference. At best, they will serve as guideposts for further research.

IV

A MULTI-FACTOR MODEL OF TIME-SERIES OF EDUCATION-INCOME RELATIONSHIPS

As mentioned earlier, rates of return on educational investment are determined by the interaction of several supply and demand effects. Such effects are likely to operate differentially over time on various age-education groups. We have already shown that the actual experience of particular cohorts varied sizeably from what would have been predicted from the use of a cross-section profile and a growth rate, and that attempts to characterize the effects of time shifts on the age-education-income profile on the basis of particular "base year" and "end year" relationships are likely to fail. It would have been convenient for the analyst if either of these ways of dealing with the effects of shifts in demand and supply relationships had proved to be adequate; one of the attractions of the simple rate of return measure is that it allows one to avoid this tangle of inter-relationships. Unfortunately, we cannot avoid dealing more explicitly with the multiplicity of economic forces affecting education-income relationships.

The major obstacle to any efforts to deal more explicitly with the changing supply and demand effects has been the lack of adequate data. Those used in the previous section are severely limited, but they are better than any other which have been available until now. On that basis, we have attempted to construct a model which, while constrained by the data limitations, will at least be illustrative of the kind of effort which needs to be undertaken.

The data which we have at hand are:

1) Mean income for males by age and education level for six years in the period 1956 to 1966 (See Table III for source).

2) The number of males by age and education level in the labour force in 1960. (From the United States Bureau of the Census 1/1000 sample data tape.)

3) National unemployment rates for the period 1956-1966.

Some preliminary results of our attempt to develop and estimate a model based on these data only are reported in the following pages.

Appendix B develops in detail the rationale for the simultaneous supply and demand equation model of the determination of income for a particular age and education group. Here we shall only discuss the reduced form equation of that system. (Of course, one may take a less structured view of this work and simply look at the final single equation which is estimated as an ad hoc construct. Given the limitations of the data and the resultant left out variables and simplified equation form, this would seem a reasonable view. The more elaborate structure in Appendix B is primarily intended to indicate directions further work might take.)

The equation to be estimated is as follows:

$$Y^{a}_{16,t} - Y^{a}_{12,t} = A^{a} + B^{a}_{1} S^{a}_{16,t} + B^{a}_{2} S^{a}_{12,t} + B^{a}_{3} S^{\Sigma a}_{16,t} + B^{a}_{4} S^{\Sigma a}_{12} + B^{a}_{5} U_{t} + B^{a}_{6} T_{t} + e_{t}$$

where:

$Y^{a}_{16,t}$ = mean annual income of males with 16 or more years of education, age a in year t,

$Y^{a}_{12,t}$ = mean annual income of males with 12 years of education, age a in year t,

$S^{a}_{16,t}$ = number of males with 16 or more years of education in the labour force, age a in year t,

$S^{a}_{12,t}$ = number of males with 12 years of education in the labour force, age a in year t,

$S^{\Sigma a}_{16,t}$ = the sum of the number of males with 16 or more years of education in the labour force with age a-2, a-1, a+1, a+2 in year t,

$S^{\Sigma a}_{12,t}$ = the sum of the number of males with 12 years of education in the labour force with age a-2, a-1, a+1, a+2 in year t,

U_{t} = the national unemployment rate in year t,

T_{t} = the number of years elapsed since 1956,

e_{t} = an error term.

The dependent variable is, for a given age, the difference between the income of a male with 16 or more years of education and of one with 12 years of education. As noted above, this is the difference which has usually been viewed as the resultant of the investment in the additional years of education. Note that there would be a different equation estimated for each age level. In this way, one would allow for a completely different impact of the various independent variables for each age level.

The four "S" independent variables are included to represent the relative supplies of different types of educated manpower. Including the S variables, made up of the sum of age groups two years on either side of the age group for the equation, represents a crude attempt to take into account the possibility of skilled labour of approximately the same age as a competing source of skill supply. In general, the difference in income between those with 16 and those with 12 years of education is taken to be influenced by the relative supply of these two groups in the same age bracket and of closely competing or complementary groups in near age brackets.

The unemployment rate is used as a general variable to represent the level of economic activity. Its effect on the dependent variable would represent the differential impact of changes in the level of economic activity on the two education level groups of a given age.

The time trend variable is, unfortunately, a catch-all variable for a number of factors. It could represent trends in the composition of demand for final goods due to rising incomes, technological trends in production relationships, and trends in the supplies of other factor inputs.

As stated above, there are only six years (within the period 1956-1966) for which we have observations for the dependent variable. With seven coefficients to be estimated, we would have no degrees

of freedom left. Therefore, rather than estimate a separate equation for each age level, it was necessary to make some further assumptions in order to pool the data for all ages and conserve degrees of freedom. In order to do this, we assumed that the relationship for any given B_i across all the age levels could be adequately approximated by a quadratic function. Thus, rather than estimate separate B_i^a, we could estimate a single equation for all the age levels of the following form:

$$Y^a_{16,t} - Y^a_{12,t} = A_a + B_1 S^a_{16,t} + B_2 a S^a_{16,t} + B_3 a^2 S^a_{16,t} + B_4 S^a_{12,t} + B_5 a S^a_{12,t} + B_6 a^2 S^a_{12,t}$$
$$+ B_7 S^{\Sigma a}_{16,t} + B_8 a S^{\Sigma a}_{16,t} + B_9 a^2 S^{\Sigma a}_{16,t} + B_{10} S^{\Sigma a}_{12,t} + B_{11} a S^{\Sigma a}_{12,t} + B_{12} a^2 S^{\Sigma a}_{12,t}$$
$$+ B_{13} U_t + B_{14} a U_t + B_{15} a^2 U_t + B_{16} T_t + B_{17} a T_t + B_{18} a^2 T_t + e_t$$

where a = 32, 33, ..., 57.

Even with this formulation, the impact of the supply variables, the unemployment and time trend variables can vary by age, since those terms with a and a^2 will differ for different age levels (our assumption of a quadratic form relating coefficients across age levels does force these differential effects to follow a certain degree of continuity).

This regression equation, quadratic with respect to age, was estimated in a stepwise fashion. In Table VIII, we report the results of the estimation where the stepwise procedure was cut off at the point at which multi-collinearity clearly began to appear (i. e. the corrected R^2 began to fall with the addition of more variables).

The results of this regression are difficult to evaluate directly from the table and in order to facilitate commentary we have, therefore, calculated the "net coefficients" for each independent variable at various age levels and charted them (Charts Va-e). The net coefficient for a given variable for a given age is obtained by combining the regression coefficients (for example, for age 32, the net coefficient for U_t is 4504.3 + (32) (-199.16) + $(32)^2$(2.0516) = 232.02).

We should like to remind the reader, as we proceed to comment on the regression results, that these estimates should be regarded as illustrative; these are only preliminary results and we intend to attempt further refinements in the estimation procedures.

Looking first at the net coefficients for U_t, the variable interpreted as reflecting the level of economic activity, we find that it has a positive effect on the differential between college and high school incomes up to the age of about 38. Over this range, a lower level of business activity (higher U_t) reduces high school graduates' incomes more than those of college graduates. After the age of 38, the effect on high school graduates' incomes would seem to be less than that on college graduates' incomes. If these estimates are even approximately correct, different age cohorts experiencing different levels of business activity at different points in the life cycle could thus have rather different lifetime earnings patterns. It might seem strange, at first, that for a substantial part of the age range the U_t net coefficients are negative, since one normally thinks of those with less education being the first to be laid off during periods of low economic activity. However, one must remember that these are measures of mean incomes and that, while better-educated individuals may not be laid off, the rate of increase in their incomes can be considerably lessened and, therefore, their differential gain over those with lower education decreased.

Turning to the net coefficients for T_t, one might compare these estimates with those of Miller and Ben-Porath discussed above. The coefficients of this variable would, in a sense, reflect the "pure

Table VIII. REGRESSION EQUATION FOR MULTI-FACTOR MODEL

Dependent Variable $Y^{a}_{16,t} - Y^{a}_{12,t}$

(N = 155)

INDEPENDENT VARIABLE	COEFFICIENT	T-RATIO	INDEPENDENT VARIABLE	COEFFICIENT	T-RATIO
A_{32}	-6580.3	-3.230	* $S^{a}_{16,t}$		
A_{33}	-5706.1	-3.039	* $a.S^{a}_{16,t}$		
A_{34}	-4848.8	-2.825	* $a^2S_{16,t}$		
A_{35}	-4023.2	-2.580	* $S_{12,t}^{a}$		
A_{36}	-3233.7	-2.305	* $a.S^{a}_{12,t}$		
A_{37}	-2485.5	-1.979	* $a^2S^{a}_{12,t}$		
A_{38}	-1789.3	-1.595	* $S^{\Sigma a}_{16,t}$		
A_{39}	-1143.0	-1.142	$a.S^{\Sigma a}_{16,t}$	-.09496	-.825
A_{40}	- 543.8	- .610	$a^2S^{\Sigma a}_{16,t}$	.00175	.601
A_{41}	6.7	.008	$S^{\Sigma a}_{12,t}$	7.9219	2.162
A_{42}	502.2	.690	$a.S^{\Sigma a}_{12,t}$	-.20764	-2.310
A_{43}	939.6	1.393	* $a^2S^{\Sigma a}_{12,t}$		
A_{44}	1299.8	2.039	U_t	4504.3	4.464
A_{45}	1596.6	2.607	$a.U_t$	-199.16	-4.365
A_{46}	1817.2	3.064	$a^2.U_t$	2.0516	4.067
A_{47}	1991.5	3.439	* T_t		
A_{48}	2085.3	3.703	$a.T_t$	-2.5684	-.736
A_{49}	2118.9	3.889	$a^2.T_t$	.14150	1.425
A_{50}	2061.9	3.969			
A_{51}	1926.7	4.000			
A_{52}	1741.2	3.992			
A_{53}	1477.1	3.855			
A_{54}	1172.5	3.667			
A_{55}	835.4	3.194			
A_{56}	434.2	2.091			

R^2 Corrected for D.F. = .9034

* Stepwise procedure reached highest corrected R^2 before this variable entered

economic growth effect" they sought to isolate. The advantage of these estimates over those of Miller and Ben-Porath is, first, that they are less sensitive to the choice of beginning or ending periods - since they take into account each of the six years of observation within the ten-year period - and, second, that they are an even "purer" measure to the extent that the other variables in the model have taken into account other effects. In spite of these advantages, one would be reluctant to call these coefficients "the effects of economic growth", since they could be due to changes in other factor supplies, tastes, or technology, as well as to the general growth in incomes. These estimates of time trend factors suggest that they cause the differential of the older groups to grow faster than that of the younger groups.

The coefficients of the supply variables, $S^{\Sigma a}_{16,\,t}$, $S^{\Sigma a}_{12,t}$, are the most difficult to interpret. Appendix B provides a more detailed technical discussion about the expected signs of these coefficients when they are regarded as the coefficients of the reduced form of the demand and supply model spelled out there. We shall make only a few points here, and leave the rest of the discussion for the Appendix. The first problem that one might note is that the $S^{a}_{16,t}$ and $S^{a}_{12,t}$ variables did not enter the equation before the stepwise procedure reached the highest corrected R^2, indicating that they were collinear with their respective Σa counterparts and that, therefore, their separate effects on the dependent variable could not be estimated. Thus, rather than four supply factors, we end up with two supply factors, i.e. our attempt to separate out the effects of those with the same education level but slightly older or younger did not work. Our idea was that these groups might be either competitive with, or complementary to, the age groups concerned, but the data at hand do not allow us to isolate such group effects. The theoretical meaning of these Σa groups, therefore, becomes rather cloudy and the results which emerge are not such as to help us very much in interpreting their meaning. Had the signs of the $S^{\Sigma a}_{16,t}$ been uniformly negative and those of $S^{\Sigma a}_{12,t}$ uniformly positive, we would have perhaps been able to interpret these factors broadly as representing two competitive supply sources. The fact that both are negative over a wide range of ages and that they each have positive coefficients at slightly different points is confusing. Furthermore, it is hard to conceive of a theoretical explanation for the strong trends in these coefficients with respect to age.

In general, the most disturbing result, from a theoretical point of view, is the fact that the coefficient of $S^{\Sigma a}_{12,t}$ is negative over a wide age range. If this variable is thought of as representing the supply of those with 12 years of education, one would expect the increase in that supply to lower $Y^{a}_{12,t}$ and to raise $Y^{a}_{16,t}$, and thus produce a positive coefficient for the variable. The inadequacies of this model are most clearly indicated by this result.

Once we have allowed simultaneously for the effects of the level of business activity, time trends and supply factors, we have a set of coefficients, the A_a, which provides us with an age-income profile for the differential associated with investment in education beyond the 12th year. We should note, first of all, that the stepwise results indicate that the age variables alone can account for about 75% of the variation in the dependent variable (the other variables add about 16 percentage points more to the explained variation). There is no question but that age is an important variable affecting education-income relationships.

Chart Ve gives the age-income differential profile provided by the model estimates. This profile was constructed by setting the unemployment and size variable at their mean values for the 1956-1966 period and T at the mid-point value, 5.5, and then calculating the income differential values for each age level.

Chart Ve also provides the profile drawn from the cross-section data for 1956, the base year. Note that neither profile allows for time trend shifts in relative incomes. The model profile is somewhat higher at the outset, peaks earlier and drops off more sharply. From casual inspection, one

Chart V a

NET COEFFICIENTS FOR U_t

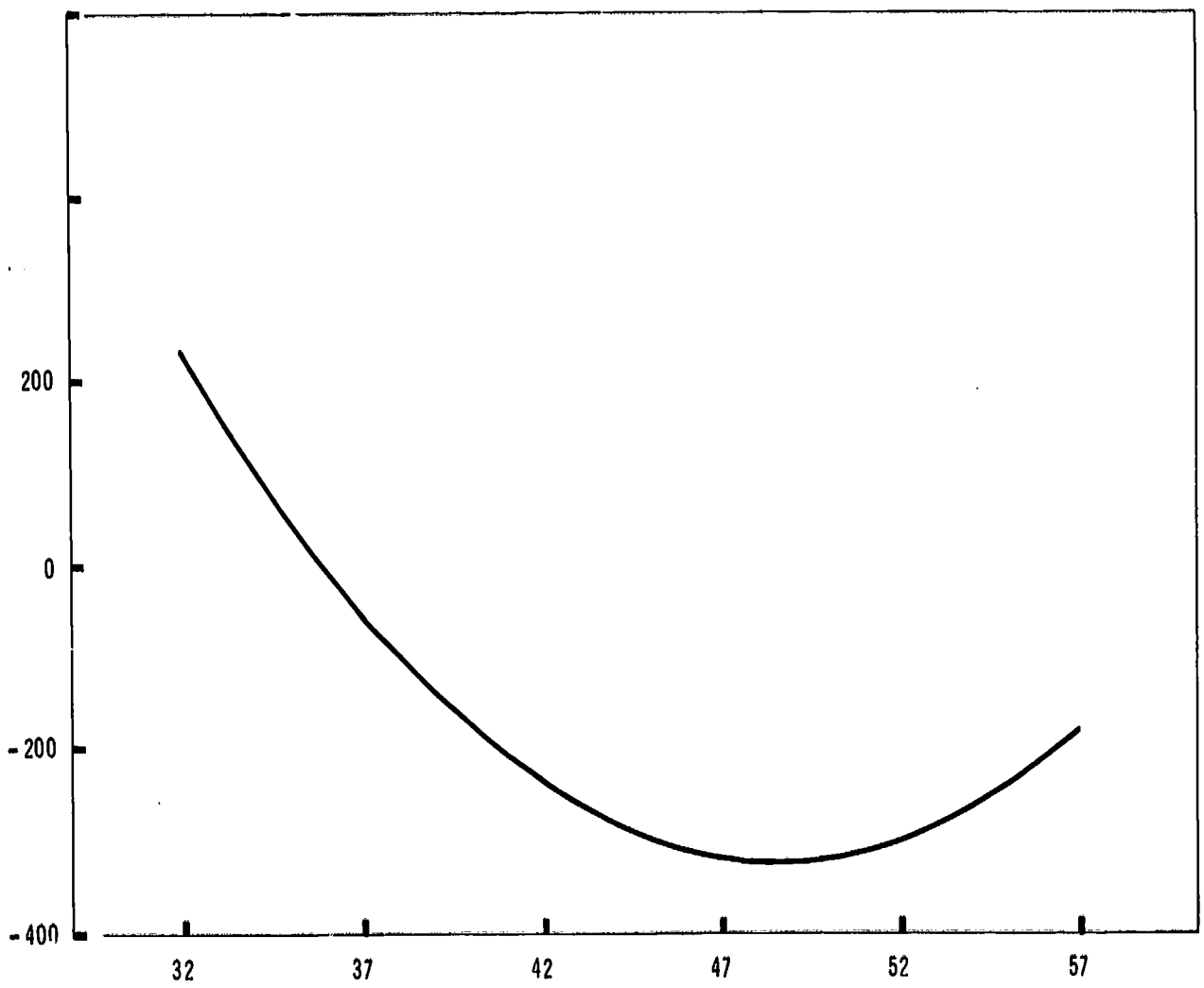


Chart Vb

NET COEFFICIENTS FOR T_t

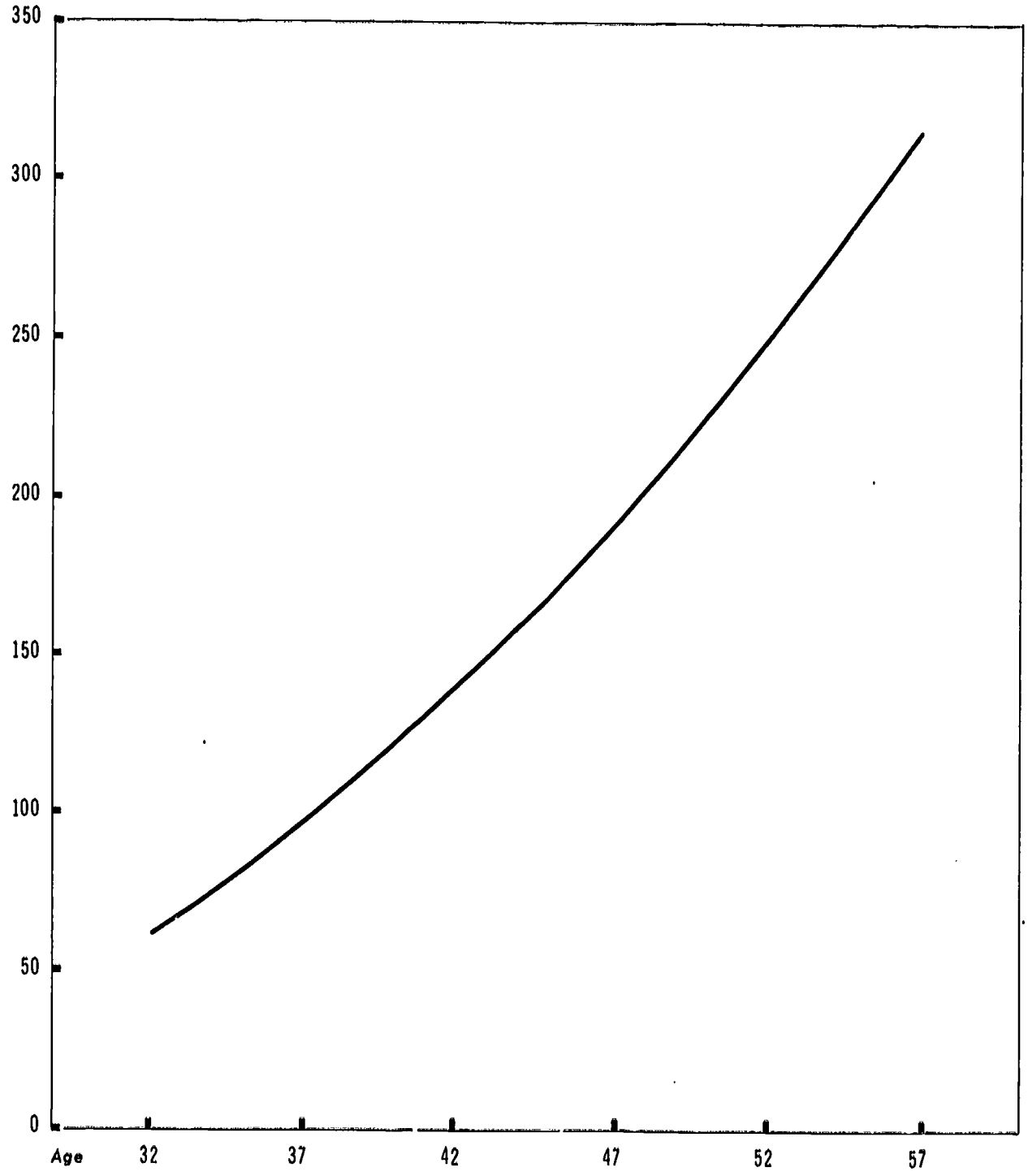


Chart Vc

NET COEFFICIENTS FOR $S_{12,t}^{\Sigma a}$

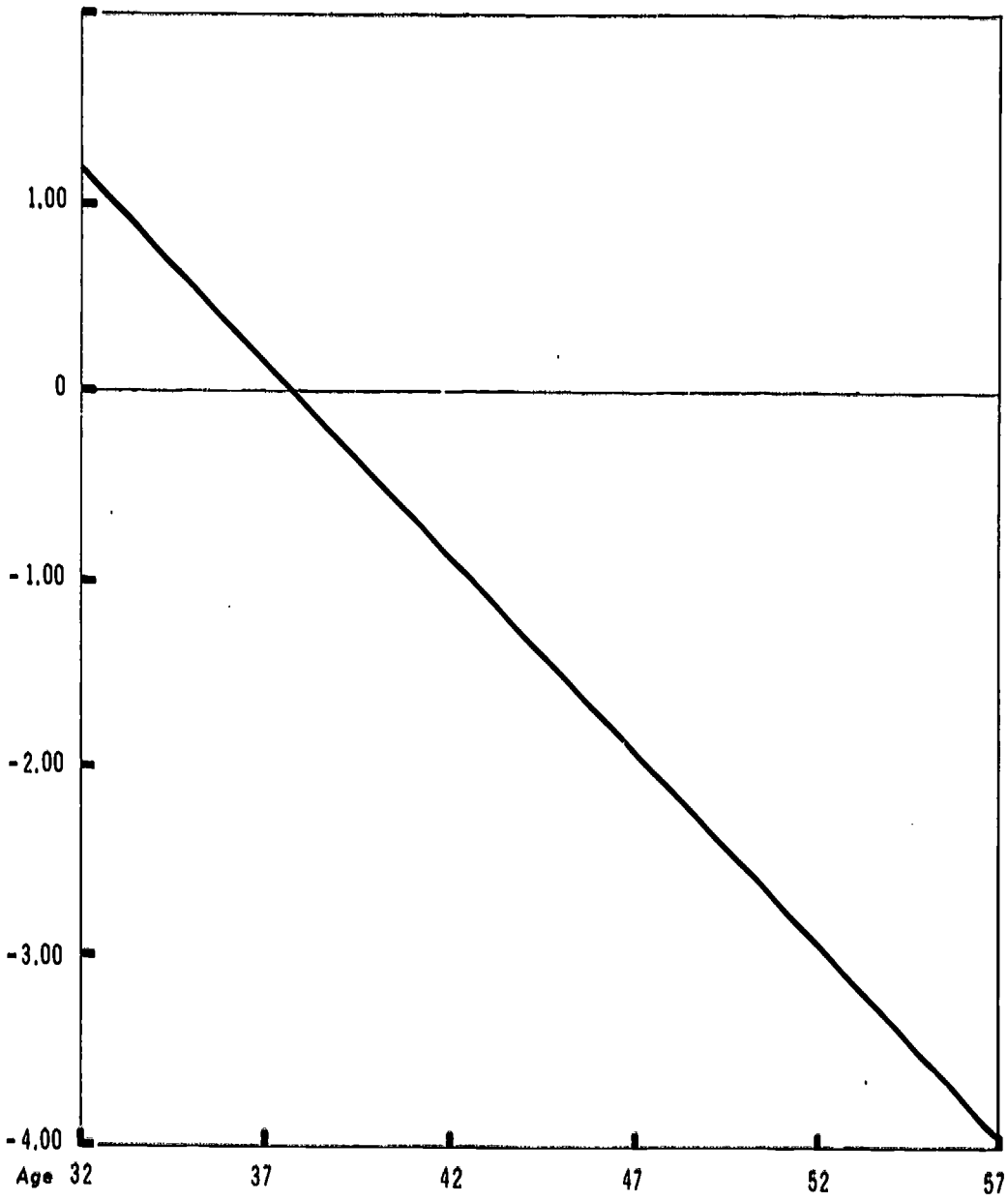


Chart V d

NET COEFFICIENTS FOR $S_{16,t}^{\Sigma a}$

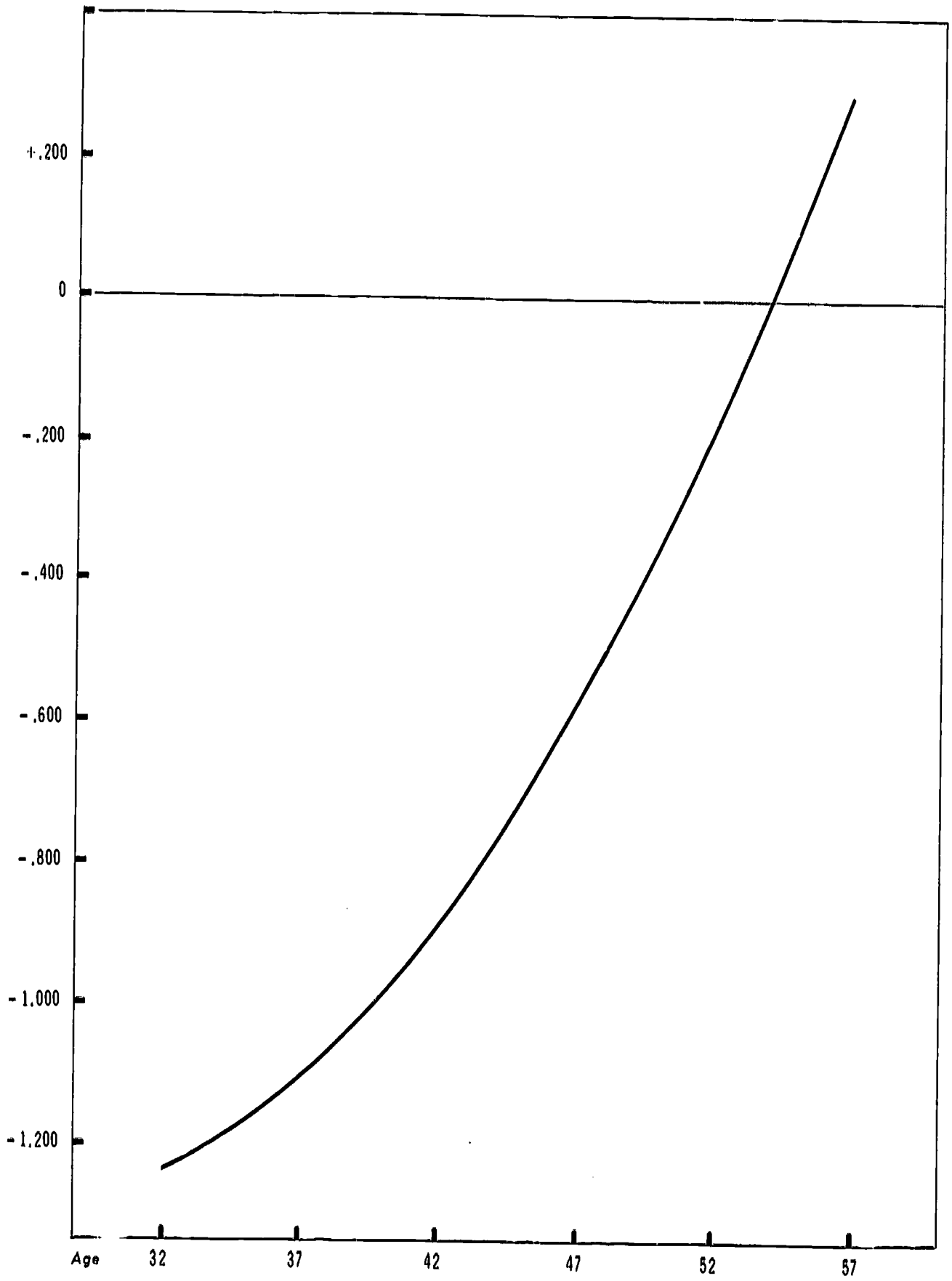
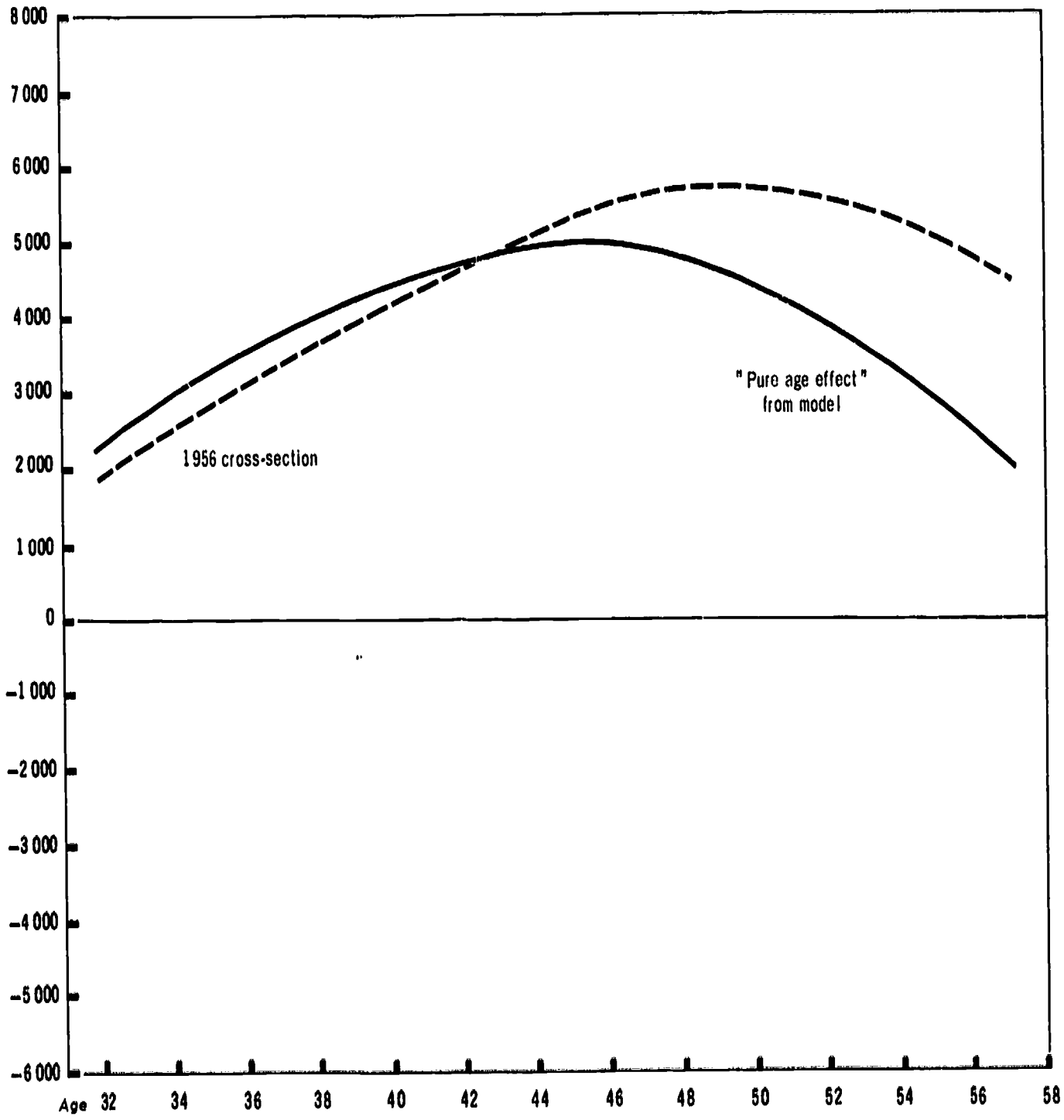


Chart Ve

AGE-INCOME DIFFERENTIAL PROFILES
(Mean income males with 16 or more years of education
minus mean income males with 12 years of education)



might conclude that at reasonable discount rates the present value of the model profile might be somewhat lower than that of the 1956 profile. However, it is important to recall that any cross-section data will include the simultaneous interaction of a number of the factors which we have sought to represent in this model, and thus one cannot decide, by looking at Chart Ve, whether the actual cross-section profile or the estimated profile comes closer to the type of measure which is basically implied by human capital theories.

There are still further implications for human capital theory which could be drawn from the particular estimates provided by this model. However, since the model was presented primarily in order to illustrate an approach that seemed fruitful, we shall forego further speculation about the meaning of these results.

V

SUMMARY AND CONCLUSIONS

The past decade has witnessed a tremendous growth in the literature devoted to the role of education in the determination of incomes and in the growth of the economy. Studies on this subject have centred around empirical estimates of income differences due to differences in the level of education. These estimates have been based almost exclusively on age-education-income profiles derived from cross-section data. The aim of this paper was to seek to answer a simple, but clearly very significant, empirical question: to what extent do the income differentials associated with education and observed at a given point in time provide a good basis for predicting the actual income differentials which will be experienced by a cohort group over time?

It would have been very convenient if it had turned out that the cross-section data did provide a simple means of predicting the future differentials of the various cohorts. However, such a finding would have seemed very strange from the point of view of traditional economic theory, since these differentials represent the relative prices of certain factors of production and the theory would lead us to believe that over time, as relative demand and supply forces shift, relative prices would also shift.

Though the data are very scarce, an attempt was made to compare income differentials for various education groups in the United States as predicted from the cross-section data at the beginning of the period, and adjusted by a growth rate, with income differentials actually experienced by various cohort groups over the period. The data used were those for a twenty-year period (for which there were three points of observation for ten-year age cohort groups) and for a ten-year period (for which the data were richer, having both 6 years of observation and single-year age cohort groups). The results showed that the cross-section data did not provide an adequate and simple basis for predicting time-series income differentials. For both the twenty-year period and the ten-year period, the actual income differentials of various cohorts differed sizeably from what would have been predicted from the cross-section data. For example, illustrative calculations suggest that actual rates of return on investment in education of some cohorts might have been as much as 6 percentage points below those predicted from the cross-section data in some cases (e.g. if the expected rate of return was 10%, the actual rate of return would be only 4%), and as much as 4.5 percentage points above what would have been predicted for other cohorts. There was no clear pattern to these differences; some cohorts exceeded, some fell short of predicted values.

It might have been hoped that these differences merely reflected some sort of simple long-term trend associated with the process of economic growth. In fact, there have been one or two comments in the literature suggesting that this was so. If this had been the case, it would still have been possible to utilize the cross-section data for prediction purposes by adding some simple trend adjustments. However, it has been shown in this paper that simple long-term trends which might be labelled "effects of economic growth" are not an adequate explanation of the changes in income differentials over time; the effects of "trends" in income differentials are shown to vary according to the choice of base years and final years over which to measure such trends. Thus, the hope that simple trend corrections of the cross-section data will provide an adequate basis for prediction must be abandoned.

The easier routes having proved fruitless, an attempt was made to formulate more complex models of the determinants of income differentials of various education groups over time. The idea that, over time, many economic and social factors are likely to interact in determining rates of return on investment in education has certainly been advanced by many of the students of human capital theory, but little has been done empirically to isolate such factors and to estimate their effects. Probably the primary reason for the absence of this sort of empirical investigation has been the lack of adequate time-series information on many of the relevant variables, most particularly age-education-income data.

Data limitations remain a serious obstacle, but the last part of this paper describes a preliminary attempt to utilize some ten-year data for the United States to estimate a crude multi-factor model of the determination of age-education-income relationships. The results of this attempt are described not because they are conclusive, but, primarily, because they serve to illustrate the kind of investigation which might have to be undertaken if a deeper understanding of age-education-income relationships is to be achieved. The model described allows for the simultaneous effects on income differences of variables taken to represent age, the level of business activity, supplies of various types of educated manpower, and time trend factors.

The following conclusions can be drawn from the results:

a) The impact of the various forces on income differentials varies substantially with age.

b) Variations in the level of business activity have substantial effects on income differentials, these effects being different at different ages (since most cross-section studies in the United States have been drawn from censuses taken at periods of relatively high unemployment, this could have had serious effects on returns estimates).

c) Time trend factors seem to be increasing differentials, faster at the older age levels.

d) Variables which were meant to reflect the relative supplies of educated manpower did not operate very well in the model, suggesting that the specification of such variables was inadequate.

It must be recalled that the model described is crude, the results preliminary, and the data limitations severe; but the results seem to indicate that this type of approach to the empirical problem is worthy of further investigation.

It might be useful to comment on how the results presented in this paper reflect on the long-standing controversy over the relative merits of the manpower planning approach and the rate of return approach to educational planning (see for example Blaug[1]). Earlier, this controversy was largely carried out in terms of the theoretical strengths and weaknesses of each approach. More recently, empirical evidence has been presented which substantiates the weaknesses of the manpower planning approach procedures utilized to date (see, for example, Bowles[2], Chapter V, Psacharopoulos[3] and Hollister[4]). Since the rate of return approach to educational planning has made use of cross-section data as the basis for estimates of returns, it must be concluded, in the light of our results, that empirical evidence substantiates the weaknesses of that approach too. In the case of both approaches, the weaknesses are the result of simplifications aimed at obtaining in the short term empirical results of some operational usefulness. In the light of our results and those of the studies which illustrate the weaknesses of the manpower planning approach, it seems clear that an adequate understanding of the relationship between education and income differentials and of the role of education in economic growth will only be reached by more intensive investigations along the lines suggested by the multi-factor model described in this paper. It has

1. Blaug, Mark, "Approaches to Educational Planning", Economic Journal, June, 1967.
2. Bowles, Samuel, Planning Educational Systems for Economic Growth, Cambridge: Harvard University Press 1969, Chapter V.
3. Psacharopoulos, George, "Estimating Shadow Rates of Return to Investment in Education", Journal of Human Resources, Winter 1970.
4. Hollister, Robinson G., A Technical Evaluation of the First Stage of the Mediterranean Regional Project, OECD, 1967.

been useful, perhaps, to ignore these complications during the last decade, while the basic elements of human capital theory were being developed. If, however, this type of work is seriously to influence future policy, these simplified procedures will no longer be adequate. Economists have long recognized that the theory of capital is a complex subject, and a difficult one to translate into operational terms. The results reported in this paper suggest that human capital theory shares these characteristics.

Appendix A

A CRUDE METHOD FOR TRANSLATING RATIOS OF PRESENT VALUES INTO RATIOS OF INTERNAL RATES OF RETURN

It may be recalled that in the application of capital theory to investment in education, either of two approaches (equivalent, except in a few unusual cases) is used:

a) the income differential stream attributed to the educational investment is reduced to a present value by discounting by the appropriate social (or private) discount rate; total costs (direct and foregone earnings) are calculated and similarly discounted, and the ratio of the present value of income differentials and the present value of costs is used as the measure of the value of the investment - as compared to benefit-cost ratios for other investments;

b) the internal rate of return is calculated by finding the rate of discount which will cause the present value of the stream of income differentials just to equal the present value of the total costs.

Tables I and II in the text show the ratio of present values of income differential streams (discounted at 6%). Since some people prefer to discuss these matters in terms of internal rates of return, we need a simple method to translate these ratios into differences of internal rates of return. For example, if the present value of the actual income differential stream is less than that of the expected stream, the actual internal rate of return will be smaller than the expected one. In general, it is not possible to make a simple translation from ratios of present values to ratios of internal rates of return, since the relationship between the two depends on the specific time-shape of each of the income differential profiles. However, if we make the rather strong assumption that the actual and expected differential streams differ only by a constant of proportionality, it is possible to provide some illustrative relationships between the two ratios.

For the purposes of the illustrative calculations, actual total costs are not needed. The problem is evaluated from the point in time at which the investor has just completed his education and is about to enter the labour market. Thus the present value of total costs is a fixed sum. It is then assumed that the expected income differential stream was such that a discount rate of 10% was just sufficient to reduce the present value of the stream of equality with the given total costs, i.e. the expected internal rate of return was 10%. The problem then is to find the rate of discount which would just reduce the actual income differential stream to the given total cost value, i.e. to find the actual internal rate of return. To do so, we proceed as follows:

1. Assume the expected internal rate of return was 10%, i.e.:

   a) $P.V._{n,10}^{e} = Y_i^e / (1 + .10)^i = TC$

   where Y_i^e = expected income differential in year i attributed to additional education,

TC = total costs of the educational investment up to year 0,

$PV^{e}_{n,10}$ = present value of the n year expected income differential stream discounted at 10%.

b) $PV^{a}_{n,10} = Y^{a}_{i} / (1 + .10)^{i}$

where Y^{a}_{i} = actual income differential in year i attributed to additional years of education,

$PV^{a}_{n,10}$ = present value of the n year actual income differential stream discounted at 10%.

2. The annualized equivalents of these two present values are:

$$PV^{e}_{n,10} = S^{e}A_{n,10}$$

$$PV^{a}_{n,10} = S^{a}A_{n,10}$$

where S^{e}, S^{a} = the annual amount equivalents,

$A_{n,10}$ = the present value of 1 dollar per year annuity for n years at a 10% discount rate.

$$\frac{PV^{e}_{n,10}}{PV^{a}_{n,10}} = \frac{S^{e}A_{n,10}}{S^{a}A_{n,10}} = \frac{S^{e}}{S^{a}}$$

This is the point at which the assumption of proportionality of the expected and actual income streams is important, for, if $Y^{a}_{i} = Y^{e}_{i}$ (p) for all i (p being the constant of proportionality), the ratio of PV^{e} and PV^{a} will be the same at any discount rate. Thus the ratios of the present values which are in Tables I and III in the text are at a 6% discount rate but,

$$\frac{PV^{e}_{n,6}}{PV^{a}_{n,6}} = \frac{S^{e}A_{n,6}}{S^{a}A_{n,6}} = \frac{S^{e}}{S^{a}} = \frac{PV^{e}_{n,10}}{PV^{a}_{n,10}}$$

3. It has been assumed that

$$TC = PV^{e}_{n,10}$$

The figure sought is that discount rate, x, which would cause:

$$PV^{a}_{n,x} = TC$$

Substituting the annualized values, the result is:

$$S^{a}A_{n,x} = TC = S^{e}A_{n,10} \quad \text{or}$$

$$A_{n,x} = \frac{S^{e}}{S^{a}} A_{n,10}$$

4. $\frac{S^{e}}{S^{a}}$ has been shown equivalent to the ratio expected present value/actual present value which appears in Tables I and III. $A_{n,10}$ is drawn from a total of present values of 1 per annum at compound interest. Multiplying the $A_{n,10}$ from the table by $\frac{S^{e}}{S^{a}}$, the value of $A_{n,x}$ is obtained and the value of x is determined from the same present value annuity table. For example:

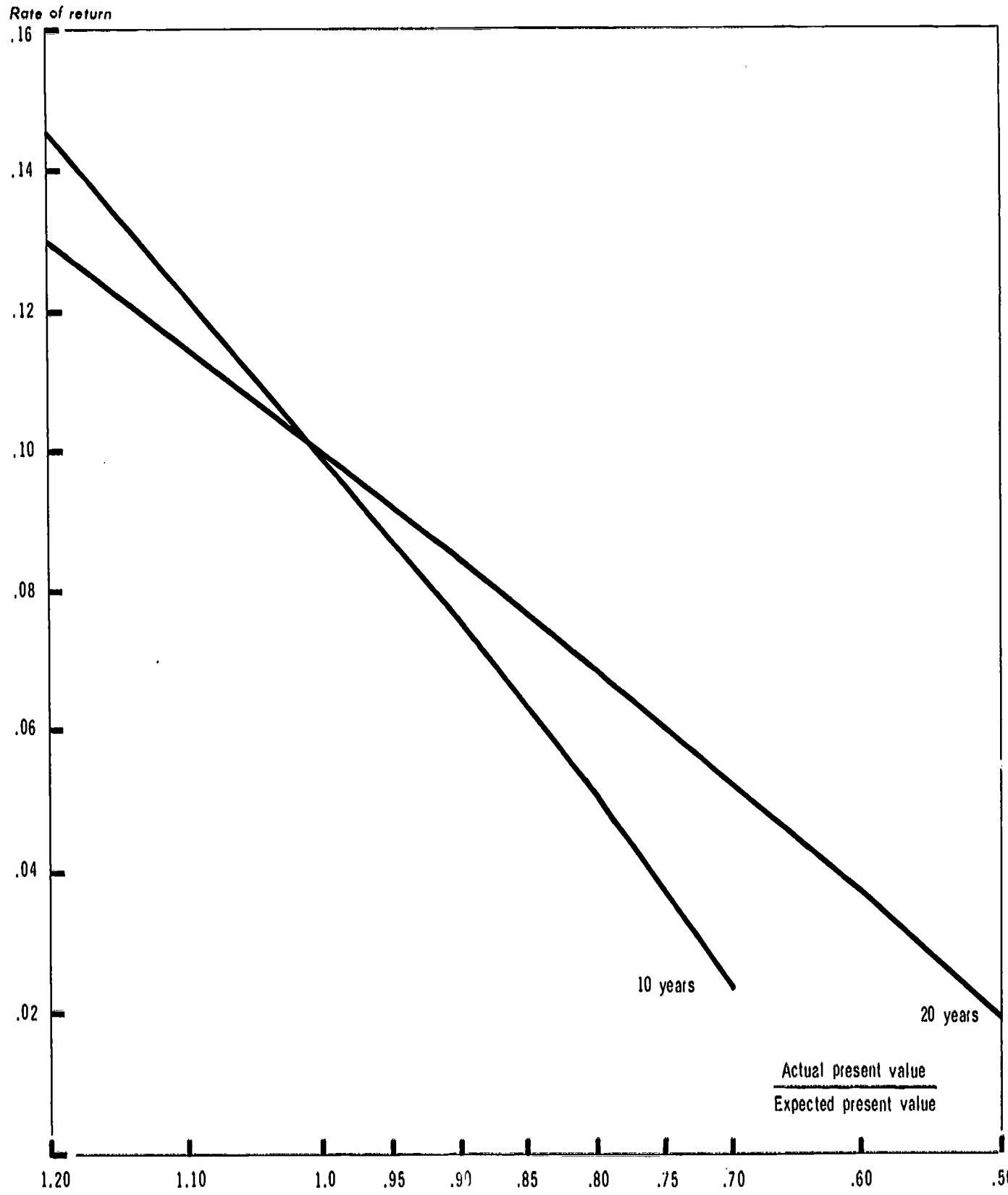
$$PV^{e}_{20,6}/PV^{a}_{20,10} = 1/.8$$

$$A_{20,10} = 8.5136$$

$A_{20,x} = 1/.8\ (8.5136) = 10.6419$ which by reference to the table yields $x = .069$.

5. The results of this translation from the ratio of present values into internal rates of return, made on the assumption that the expected internal rate of return was 10%, are provided in Chart A. The several assumptions made in this translation make it only a very rough means of approximation of relationships between relative present value and internal rates of return. The results should, therefore, be regarded as very approximate.

Chart A


Rate of return
.16
.14
.12
.10
.08
.06
.04
.02
10 years
20 years
Actual present value
Expected present value
1.20
1.10
1.0
.95
.90
.85
.80
.75
.70
.60
.50

Appendix B

A SUPPLY AND DEMAND EQUATION MODEL FOR AGE - EDUCATION - INCOME RELATIONSHIPS AND SOME EVALUATION OF ITS COEFFICIENTS

1. Define

$D^{A}_{16,t}$ = demand in year t for labour age A with 16 years or more of education,

$P^{A}_{16,t}$ = price (yearly income) in year t for labour age A with 16 years or more education,

$P^{A}_{12,t}$ = price (yearly income) in year t for labour age A with 12 years of education,

$p^{\Sigma A}_{16,t}$ = average (weighted) price (yearly income) in year t for labour age A-1, A-2, A+1, A+2 ($=\Sigma A$) with 16 years or more of education,

$P^{\Sigma A}_{12,t}$ = average (weighted) price (yearly income) in year t for labour age A-1, A-2, A+1, A+2 ($=\Sigma A$) with 12 years of education,

U_t = unemployment rate in year t,

T_t = years elapsed since 1956 in year t,

e_{it} = error term for equation i.

Then assuming a linear demand function for labour age A with 16 years or more of education:

$$D^{A}_{16,t} = a^{A}_{11} + b_{11} P^{A}_{16,t} + b_{12} P^{A}_{12,t} + b_{13} p^{\Sigma A}_{16,t} + b_{14} P^{\Sigma A}_{12,t} + b_{15} U_t + b_{16} T_t + e_{it}$$

Define:

$D^{A}_{16,t}$ = demand in year t for labour age A with 16 years or more education,

$D^{A}_{12,t}$ = demand in year t for labour age A with 12 years education,

$D^{\Sigma A}_{16,t}$ = demand in year in year t for labour age A-1, A-2, A+1, A+2 with 16 years or more education,

$D^{\Sigma A}_{12,t}$ = demand in year t for labour age A-1, A-2, A+1, A+2 with 12 years education.

Then the demand functions for such labour are:

$$D^{A}_{12,t} = a_{21}{}^{A} + b_{21} P^{A}_{16,t} + b_{22} P^{A}_{12,t} + b_{23} P^{\Sigma A}_{16,t} + b_{24} P^{\Sigma A}_{12,t} + b_{25} U_t + b_{26} T_t + e_{2t}$$

$$D^{\Sigma A}_{16,t} = a_{31}{}^{A} + b_{31} P^{A}_{16,t} + b_{32} P^{A}_{12,t} + b_{33} P^{\Sigma A}_{12,t} + b_{34} P^{\Sigma A}_{12,t} + b_{35} U_t + b_{36} T_t + e_{3t}$$

$$D^{\Sigma A}_{12,t} = a_{41}{}^{A} + b_{41} P^{A}_{16,t} + b_{42} P^{A}_{12,t} + b_{43} P^{\Sigma A}_{16,t} + b_{44} P^{\Sigma A}_{12,t} + b_{45} U_t + b_{46} T_t + e_{4t}$$

Define:

$S^{A}_{16,t}$ = supply in year t of labourers age A with 16 or more years of education,

$S^{A}_{12,t}$ = supply in year t of labourers age A with 12 years of education,

$S^{\Sigma A}_{16,t}$ = supply in year t of labourers age A-1, A-2, A+1, A+2 with 16 years of education,

$S^{\Sigma A}_{12,t}$ = supply in year t of labourers age A-1, A-2, A+1, A+2 with 12 years of education.

It is assumed that supply of labour is taken as constant for any year, so that the supply functions are simply:

$$S^{A}_{16,t} = S^{A'}_{16,t}$$

$$S^{A}_{12,t} = S^{A'}_{12,t}$$

$$S^{\Sigma A}_{16,t} = S^{\Sigma A'}_{16,t}$$

$$S^{\Sigma A}_{12,t} = S^{\Sigma A'}_{12,t}$$

where S' is the fixed supply in a given year.

With these eight supply and demand equations for each age group, and the equilibrium conditions for all types of labour where D = S, it is possible to calculate the price of each type of skilled labour for each age group.

The price equations are:

$$P^A_{16,t} = C^A_1 + d^A_{11} S^A_{16,t} + d^A_{12} S^A_{12,t} + d^A_{13} S^{\Sigma A}_{16,t} + d^A_{14} S^{\Sigma A}_{12,t} + d^A_{15}$$

$$U_t + d^a_{16} T_t + W_{1t}$$

$$P^A_{12,t} = C^A_2 + d_{21} S^A_{16,t} + d^A_{22} S^A_{12,t} + d^A_{23} S^{\Sigma A}_{16,t} + d^A_{24} S^{\Sigma A}_{12,t} + d^A_{25}$$

$$U_t + d^A_{26} T_t + W_2 t$$

Subtracting $P^A_{12,t}$ from $P^A_{16,t}$

$$P^A_{16,t} - P^A_{12,t} = C^A + D^A_1 S^A_{16,t} + D^A_2 S^A_{12,t} + D^A_3 S^{\Sigma A}_{16,t} + D^A_4 S^{\Sigma A}_{12,t} + D^A_5 U_t + D^A_6 T_t + W_t$$

This is the basic formulation of the text equation on page 138, except that notation is shifted by substituting P for Y, A for C, B for D, e for W and a for A.

2. If the further assumptions are made that coefficients of the reduced form $P^A_{16,t} - P^A_{12,t}$ are related among the various A by the following equations, quadratic in A.

$$D^A_1 = D_1 + D_2 A + D_3 A^2$$

$$D^A_2 = D_4 + D_5 A + D_6 A^2$$

$$D^A_3 = D_7 + D_8 A + D_9 A^2$$

$$D^A_4 = D_{10} + D_{11} A + D_{12} A^2$$

$$D^A_5 = D_{13} + D_{14} A + D_{15} A^2$$

$$D^A_6 = D_{16} + D_{17} A + D_{18} A^2$$

then the 26 $P^A_{16,t} - P^A_{12,t}$ equations for A = 32, ..., 57 can be reduced to the single equation:

$$\begin{aligned} P^A_{16,t} - P^A_{12,t} = {} & C_A + D_1 S^A_{16,t} + D_2 A S^A_{16,t} + D_3 A^3 S^A_{16,t} + D_4 S^A_{12,t} + D_5 A S^A_{12,t} \\ & + D_6 A^2 S^A_{12,t} + D_7 S^{\Sigma A}_{16,t} + D_8 A S^{\Sigma A}_{16,t} + D_9 A^2 S^{\Sigma A}_{16,t} + D_{10} S^{\Sigma A}_{12,t} \\ & + D_{11} S^{\Sigma A}_{12,t} + D_{12} A^2 S^{\Sigma A}_{12,t} + D_{13} U_t + D_{14} A U_t + D_{15} A^2 U_t \\ & + D_{16} T_t + D_{17} A T_t + D_{18} A^2 T_t + W_t, \end{aligned}$$

which is the formulation of the text equation on page 139, except that the notation is shifted by substituting P for Y, A for C, B for D, e for W, and a for A. This is the equation which was estimated.

3. It is useful to try to deduce the expected sign of the reduced form coefficients using some assumptions about the sign and relative size of the coefficients of the structural demand equations.

   a) Although the basic model is formulated with four supply factors, $S^{A}_{16,t'}$ $S^{A}_{12,t'}$ $S^{\Sigma A}_{16,t'}$ $S^{\Sigma A}_{12,t'}$ it is useful, first, to try to evaluate the sign of reduced form coefficients for a model with only two supply factors. This step is useful because the reduced form coefficients are easier to manipulate with just two supply factors, but, more important, because in the actual estimation process $S^{A}_{16,t}$ and $S^{A}_{12,t'}$ were so collinear with $S^{\Sigma A}_{16,t}$ and $S^{\Sigma A}_{12,t}$ that they would not be distinguished statistically. Thus, in the estimation sense, the model was reduced to a two supply factor model. The equations of the two supply factor model would be:

$$D^{A}_{16,t} = a^{A}_{1} + b_{11}\, P^{A}_{16,t} + b_{12}\, P^{A}_{12,t} + b_{13}\, U_t + b_{14}\, T_t + e_1 t$$

$$D^{A}_{12,t} = a^{A}_{2} + b_{21}\, P^{A}_{16,t} + b_{22}\, P^{A}_{12,t} + b_{23}\, U_t + b_{24}\, T_t + e_2 t$$

$$S^{A}_{16,t} = S^{A}_{16,t}$$

$$S^{A}_{12,t} = S^{A'}_{12,t}$$

The reduced form of the equations $P^{A}_{16,t}$ minus the reduced form equation for $P^{A}_{12,t}$ would be:

$$P^{A}_{16,t} - P^{A}_{12,t} = C^{A} + D^{A}_{1}\, S^{A}_{16,t} + D^{A}_{2}\, S^{A}_{12,t} + D^{A}_{3}\, U_t + D^{A}_{4}\, T_t + W_t$$

the coefficients to be evaluated being D^{A}_{1} and D^{A}_{2}.

From the solution of the structural equations, the expressions for the coefficients are obtained:

$$D^{A}_{1} = \frac{1}{b_{11} - \frac{b_{12}\, b_{21}}{b_{22}}} + \left(\frac{b_{21}}{b_{22}}\right)\left(\frac{1}{b_{22} - \frac{b_{12}\, b_{21}}{b_{11}}}\right) = \frac{b_{22}}{b_{11}\, b_{22} - b_{12}\, b_{21}} + \frac{b_{21}}{b_{11}\, b_{22} - b_{12}\, b_{21}}$$

$$D^{A}_{2} = \left(\frac{b_{12}}{b_{22}}\right)\left(\frac{1}{b_{11} - \frac{b_{12}\, b_{21}}{b_{22}}}\right) = \frac{1}{b_{22} - \frac{b_{12}\, b_{21}}{b_{11}}} = \frac{-b_{12}}{b_{11}\, b_{22} - b_{12}\, b_{21}} - \frac{b_{11}}{b_{11}\, b_{22} - b_{12}\, b_{21}}$$

   b) Recalling that the b_{ij} are the own-price response of demand (when $i = j$) and the cross-price response of demand (when $i \neq j$), the following assumptions would seem reasonable for the normal case:

assume:

b_{11}, b_{22} <0; b_{12} b_{21} >0; $1b_{11}$ 1> $1b_{12}$ 1; $1b_{22}$ 1 >$1b_{21}$ 1

The denominators of the terms in both equations are positive, since:

b_{11} b_{22} >0, b_{12} b_{21} >0 and b_{11} b_{22} > b_{12} b_{21}.

Looking first at D_1^A, b_{22} is negative and greater than the positive numerator b_{21}, so the sign of D_1^A is negative. For D_2^A the numerator ($-b_{11}$) is positive and greater than the negative numerator ($-b_{12}$), so the sign of D_2^A is positive.

Thus, if the two supply factor is the relevant one (since the other two factors are statistically inseparable), the expectation should be that the sign of the D_1^A coefficient is negative and the sign of the D_2^A is positive. This expectation is not satisfied for all ages (values of A) in the estimated model.

4. When the coefficients of the reduced form of the four supply factor model are subject to an attempt at a similar evaluation, the result is not so clear. The reduced form coefficient is a quite complex combination of sums and cross-products of the structural b_{ij} and does not yield, with manipulation, a form which can be clearly evaluated in a fashion similar to the two factor supply model. The signs of the coefficient seem to depend on the relative size as well as the sign of the b_{ij}, and simple assumptions, such as those used in 3 b), do not yield an unambiguous outcome. It does not seem possible, therefore, to develop clear expectations about the signs of the D_1^A, ..., D_n^A in the four supply factor case. Thus, if the correct model is a four supply factor model, positive or negative signs of the supply coefficients would not lead to an immediate concern about misspecification of the model.

Paper 4

PATTERNS OF RATES OF RETURN TO INVESTMENT IN EDUCATION: SOME INTERNATIONAL COMPARISONS

by

W. Lee Hansen[1]

1. W. Lee Hansen is Professor of Economics and Educational Policy Studies, University of Wisconsin, and Research Associate, Institute for Research on Poverty. The research reported here was supported by funds granted to the Institute for Research on Poverty at the University of Wisconsin by the Office of Economic Opportunity, pursuant to the provisions of the Economic Opportunity Act of 1964.

The research assistance of Paul Christensen and the comments of Robinson Hollister and Louis Emmerij are gratefully acknowledged.

The financial support of the Ford Foundation is gratefully acknowledged.

GENERAL INTRODUCTION

The paper reviews the numerous studies for different countries which report rates of return to investment in different levels and amounts of schooling. An effort is made to assess their comparability, to determine whether any empirical generalizations can be derived, to explore the general nature of the policy conclusions drawn, and to suggest some of the directions that future work on rate of return patterns should take. A special effort, though not a very successful one, is made to relate various oducational distribution data to the observed rate of return patterns.

I

INTRODUCTION

A dominant theme in the American work on the economics of education is the rate of return approach to decisions about human resource allocation. This stands in marked contrast to much of the European work and to one stream of American work which reflects a planning approach. Not only do these two approaches differ, but they also indicate in their purest forms rather divergent ways of viewing the labour market and the education-training market. The purpose of this paper is not to fan the flames of controversy by arguing the superiority of one approach over the other - more likely they complement each other, as has been suggested by Blaug[1]. Instead, this paper reviews the now numerous rate of return studies, to determine whether any empirical generalizations can be derived from them, to explore the general nature of the policy conclusions which have been drawn from them, and to suggest the directions that future work in this area should take. In doing so, a special effort has been made to examine the existing educational distribution data - educational attainment, school enrolment patterns, and the like - to help explain the observed rate of return patterns. This effort, while only moderately successful, did produce a clearer idea of what other types of education distribution data are needed and what other analyses are essential to interpreting the varying patterns in the rates of return to educational investment in different countries. Hence, this paper represents an initial foray into an uncharted area and thus is far from being a definitive piece of work.

THE RATE OF RETURN APPROACH

The rate of return approach, developed largely by Becker[2,3] and Schultz[4,5], characterizes much of the initial work on the economics of human investment in the United States during the late 1950's and early 1960's. In essence, this approach recognizes that human investments in education involve cost outlays - to the individual and to society - which are expected to produce a stream of benefits, largely in the form of higher earnings over the working life of those who acquire schooling. The internal rate of return summarizes in a convenient way the relationship between the costs which are concentrated over a short span of years and the benefits which accrue over a much longer and more distant time horizon. The mechanics of the calculations and the definitions of costs and benefits ordinarily employed, as well as important qualifications to rate of return studies, need not be reviewed here (Becker[2], Hansen[6]).

1. Blaug, M., "Approaches to Educational Planning", Economic Journal, June 1967, pp. 262-288.
2. Becker, G.S., "Underinvestment in College Education", American Economic Review, May 1960, pp. 346-354.
3. Becker, G.S., Human Capital, Princeton: Princeton University Press, 1964.
4. Schultz, T.W., "Education and Economic Growth", in N.B. Henry ed., Social Forces Influencing American Education, Chicago: University of Chicago Press, 1961, Part II, pp. 46-88.
5. Schultz, T.W., "The Rate of Return in Allocating Investment Resources to Education", Journal of Human Resources, Summer 1967, pp. 293-310.
6. Hansen, W.L., "Comment on Harberger's Estimate of Rate of Return to Investment in Education in India", Mimeographed, August 7, 1963.

The usefulness of the rate of return approach to questions about human resource allocation has been stressed by Schultz[1], Johnson[2], and Solow[3] among others[4]. All of them emphasize the need for a broad concept of capital, one that embraces the major stocks of productive resources - physical and human alike. They also maintain that only by knowing the relative returns to those differing inputs can effective decisions be made about the whole gamut of activities and programmes involving education and training

While a major concern of the United States has been that of achieving greater economic efficiency, in the narrower sense, that is, making the best use of existing resource inputs, many nations are concerned with finding efficient ways to greatly and quickly augment the quantity and quality of their human resource inputs. The need for a better educated and trained labour force has long been apparent. But in the 1950's this need was dramatized by the studies of Fabricant[5] and Kendrick[6] who discovered the "residual" - the large increment to economic growth left unexplained by conventional labour and physical capital inputs - and by Denison[7] whose pathbreaking work attributed a good part of the residual to education and the production of new knowledge. As a consequence of these efforts, there appears to be a growing convergence of interest in rate of return analyses and contribution-to-growth studies in the less-developed countries (Gounden[8,9]; Williamson[10,11]). Contribution-to-growth studies indicate the role of education in accounting for past economic growth and are also suggestive of the effects on future economic growth. Rate of return studies complement growth studies through their focus on the various levels of schooling. Thus, they show more precisely the relationship between the benefits and costs of different types and levels of schooling, in the recent past and presumably in the near future as well.

As yet we know little about the relationship, if any, between the rates of return to the various levels of schooling and other characteristics of the economy and society - its level and rate of development, the level and distribution of educational attainment, the current flow of graduates from the educational system, and the like. Carnoy[12] has speculated on some of these relationships, and Harbison and Myers[13] have attempted some analyses along these lines, though without reference to rates of return to schooling.

1. Schultz, T.W., "The Rate of Return in Allocating Investment Resources to Education", Journal of Human Resources, Summer 1967, pp. 293-310.
2. Johnson, H.G., "Comment" in The Residual Factor in Economic Growth, Organisation for Economic Cooperation and Development, Paris 1964, pp. 219-227.
3. Solow, R.M., Capital Theory and the Rate of Return, Amsterdam: Rand McNally, 1963.
4. The criticisms of the rate of return approach are numerous and no attempt is made here to cover that well-explored territory; for a good review of the entire discussion, see: Pandit, H.N., Productivity of Investment in Education: the Conceptual Framework, Techniques and Major Findings, New Delhi: National Institute of Education, 1969.
5. Fabricant, S., Basic Facts on Productivity Change, New York: National Bureau of Economic Research, 1959.
6. Kendrick, J.W., Productivity Change in the U.S., Princeton: Princeton University Press, 1961.
7. Denison, E.E., The Sources of Growth in the U.S. and the Alternatives Before Us, Washington, D.C.: Committee for Economic Development, 1962.
8. Gounden, A.M.N., "Education and Economic Development", unpublished Ph. D. dissertation, Karukshetra University, India, 1965.
9. Gounden, A.M.N., "Investment in Education in India", Journal of Human Resources, Summer 1967, pp. 347-358.
10. Williamson, J.G., "Economic Growth in the Philippines: 1947-1965: The Role of Traditional Inputs, Education, and Technical Change", Institute of Economic Development and Research, School of Economics, University of the Philippines, Discussion Paper No. 67-8, September 12, 1967.
11. Williamson, J.G., and DeVoretz, D.J., "Education as an Asset in the Philippine Economy", in M.B. Conception, ed., Philippine Population in the Seventies, Manila: Community Publishers, Inc., 1969, pp. 133-168.
12. Carnoy, M., "Rates of Return to Schooling in Latin America", Journal of Human Resources, Summer 1967, pp. 359-374.
13. Harbison, F. and Myers, C.A., Education, Manpower and Economic Growth, New York: McGraw-Hill, 1964.

A comparative examination of the available rate of return studies now seems appropriate in order to determine what broader generalizations, if any, can be drawn from them[1]. Part II describes the studies which are available, their temporal and geographic coverage, statistical base and representativeness, and then reviews some of the major methodological problems encountered in comparing these studies. Part III compares the empirical results, and Part IV indicates what conclusions can be reached and what the priorities should be for future research on this topic.

1. No effort is made to examine "shadow rates of return" generated through linear programming models, see: Psacharopoulos, G., "Estimating Shadow Rates of Return to Investment in Education", Journal of Human Resources, Winter 1970, pp. 34-50.

II

SAMPLE OF STUDIES

During the past few years I have accumulated approximately twenty rate of return studies. While not systematic, the effort to collect these studies has been at least partly deliberate, with an eye to preparing a paper of this sort; undoubtedly, there are other studies which have not come to my attention. While a few of these studies are for the United States and for other developed countries, the bulk of them are for less-developed countries actively seeking to speed their rate of economic growth. Included are the following countries: Kenya, Northern Rhodesia, Uganda, India (several studies), Philippines, Israel, Great Britain, Greece, Chile (several studies), Columbia (several studies), Mexico (several studies), United States (several studies), and Canada. Appendix Table A presents a full listing of these studies, indicating author, country, year and scope of the study, as well as the data base. Most of the studies are for the late 1950's and early 1960's, although at least one extends back to the 1940's.

The extent to which these studies are comparable is not fully clear. For one thing, the coverage of political units varies considerably; though the analyses often are made at the national level, they are confined in some cases to particular geographic areas within a country. Moreover, they frequently apply to specific sectors of the economy rather than the economy as a whole, whatever the geographical coverage. The reported levels of schooling usually differ somewhat because the structure of each country's educational system varies. In addition, rates of return are not always available for certain levels of education, in particular, for literacy versus non-literacy among the non-formally educated, and for different types and amounts of post-secondary education. To further complicate matters the data base for the studies is rarely the same: some rely upon census-type data while others employ special survey data whose quality no doubt varies; some studies are based upon income and others upon earnings. Finally, the methodology for deriving the costs and return streams differs in detail even though the same general approach is usually followed.

Assessing the effects of these many differences on the comparability of the rates of return is a major task, requiring a careful evaluation of each study and ultimately a recasting of them on a comparable basis. Since such an effort would, even if undertaken, still leave a good deal of uncertainty, we shall for the purposes of this paper assume that approximate comparability exists.

COMPARABILITY

A number of more substantive problems, aside from those just mentioned, arise in comparing the results of the various studies. These problems can be grouped into four major categories: 1) use of present value versus rate of return approach; 2) use of shortcut methods to estimate benefit streams; 3) use of unadjusted data versus data adjusted for other non-school-related characteristics associated with earnings differences; 4) use of an economic growth factor in adjusting the cross-section age-earnings profiles. Each of these problems will be considered briefly. The three right-hand columns in Appendix Table B attempt to summarize for each study how these problems were handled.

1) Present Value versus Rate of Return

We shall use the rate of return criterion because most studies employ this rather than the present value net of cost. Moreover, the results are rarely reported in enough detail to permit the calculation of present figures. It is often possible, however, to infer the general magnitudes of the rates of return from present value figures, especially when the results are presented using several different discount rates.

A larger issue only touched upon here involves the underlying theoretical and empirical merits of the internal rate of return as compared to the present value. Alchian[1], Hirschleifer[2], and Bailey[3] among others, have argued for the superiority of present value over rate of return, so as to avoid multiple rates of return caused by the sometimes irregular behaviour of the age-education-earnings profiles. But in a recent note, Jean[4] demonstrates that some of the examples typically used to cast doubt upon the rate of return approach represent rather special cases; he also shows which types of age-cost-return streams yield indeterminate solutions. In most rate of return studies, the age-cost-return streams are not of the type that produce multiple solutions. In any case, however, the available data force us to concentrate on the rate of return.

2) Use of Shortcut Methods

An early problem in calculating rates of return arose because of the absence of age-earnings[5] profiles by levels of schooling. Although this problem diminished as more data became available, some of the early studies are flawed because the shortcut methods of constructing age-earnings profiles led to over or understatements of the rates of return.

One shortcut method assumes that average differences in earnings by level of schooling adequately reflect the actual pattern of differences by age level. By ignoring the fact that earnings differences tend to grow with age - after the investment period these differences becomes increasingly large with age - the effect is to increase the weight of benefits relative to costs and thereby to inflate the rate of return. The resulting overstatement of the rate of return will vary depending upon the extent to which earnings differences do increase with age. For example, Baldwin's[6] study of Northern Rhodesia uses this shortcut approach, although he builds in an offset to the overstatement that would otherwise occur. A somewhat similar approach is followed by Shoup[7] for Venezuela where he assumes constant differentials or, in some cases, builds in rather arbitrary increases in earnings with age.

The other method involves constructing synthetic age-earnings profiles, based on a variety of assumptions but relying heavily upon some observed age-earnings patterns for another region or country. Although this method could produce an over or understatement in the rates of return, the one case which has been examined - Harberger's[8] study for India - produced an understatement (Hansen[9]). Using his same assumptions on United States data for 1949, I found that, for four years of college, the synthetic data produced rates of return that ranged from one to almost two percentage points below

1. Alchian, A.A., "The Rate of Interest, Fisher's Rate of Return Over Cost, and Keynes Internal Rate of Return", American Economic Review, December 1955, pp. 938-943.
2. Hirschleifer, J.A., "On the Theory of Optimal Investment Decision", Journal of Political Economy, August 1958, pp. 329-352.
3. Bailey, M.J., "Formal Criteria for Investment Decisions", Journal of Political Economy, October 1959, pp. 476-483.
4. Jean, W.H., "Reply", Journal of Business, March 1969, pp. 99-100.
5. We shall use the term "age-earnings" even though some of the data are for "age-income" profiles.
6. Baldwin, R.E., Economic Development and Export Growth, Berkeley: University of California Press, 1966.
7. Shoup, C.S., The Fiscal System of Venezuela, Baltimore: Johns Hopkins Press, 1959.
8. Harberger, A.C., "Investment in Man Versus Investment in Machine", in C.A. Anderson and M.J. Bowman, Eds., Education and Economic Development, Chicago: Aldine, 1965, pp. 11-50.
9. Hansen, W.L., "Total and Private Rates of Return to Investment in Schooling", Journal of Political Economy, April 1963, pp. 128-141.

the actual rates of return, and, for four years of high school, the synthetic data produced rates of return between three and one-half and four and one-half percentage points below the actual rate of return. Thus, Harberger's rates of return could be understated by 8-17% for college and by 30-40% for high school. Offsetting this to some degree are the mortality effects which Harberger ignored but whose effects are less substantial. Kothari's[1] study which follows Harberger's methodology undoubtedly contains similar biases.

While the magnitude of error is likely to be smaller using synthetic profiles rather than flat profiles, i.e. those based on average differences in earnings of people with different educational attainments, we are in the somewhat more difficult position of not being certain whether the resulting rates of return are too high or too low. In any case, however, the greater plentitude of data has steadily reduced the need for constructing synthetic age-earnings profiles by level of schooling.

3) Other Adjustments

Initially, investigators had to work with whatever data were available. In the case of national or regional data, other important factors which might affect earnings differently by level of schooling could not be statistically controlled, such as family background, place of residence, etc. One way of obtaining what might be called "cleaner" estimates of the impact of schooling is to limit the sample to relatively homogeneous groups, though the cost is usually a great amount of information loss. The alternative is to use regression analysis to derive age-education earnings profiles which "hold constant" the effects of other often important independent variables. Carnoy[2],[3] experimented with several different sets of data; one was unadjusted, another was adjusted for father's occupation, industry, city of occupation, and attendance. The latter reduced the differentials in earnings attributable to schooling, since some of these other variables were correlated with schooling. Hanoch[4],[5] who carried out an even more elaborate adjustment using the abundant data from the 1960 U.S. Census 1/1,000 sample, found a similar reduction in the earnings differential attributable to schooling. Data limitations ordinarily prevent adjustments such as these, not to mention other desirable adjustments for differential ability and numerous other variables affecting incomes. Indeed, the adjustments for ability have usually been quite arbitrary, assuming that anywhere from zero to half of the observed earnings differences are due to ability and school-related factors rather than to schooling as such (Denison[6], Gounden[7],[8]). Finally, and surprisingly, many investigators have made no allowance for expected mortality, differential unemployment rates, or labour force participation patterns among groups with different amounts of schooling.

4) Economic Growth

While it is generally recognized that expected earnings will, because of economic growth, be greater than those indicated by cross-section age-earnings profiles, relatively few studies have made such adjustments. It is difficult to know what the reason is, except that of convenience; however, Hollister[9] recently suggested that cyclical fluctuations in economic activity alter age-earnings profiles differently for people with different levels of school attainment. A simple "rule-of-thumb" correction calls for adding the assumed rate of per-worker economic growth to the rate of return calculated from cross-section data. We shall have to be content with such crude corrections until the needed longitudinal data on education-age-earnings profiles become available.

1. Kothari, V.N., "Returns to Education in India", mimeographed, University of Baroda, 1966.
2. Carnoy, M., "The Cost and Return to Schooling in Mexico: A Case Study", unpublished Ph. D. dissertation, University of Chicago, 1964.
3. Carnoy, M., "Rates of Return to Schooling in Latin America", op. cit.
4. Hanoch, G., "Personal Earnings and Investment in Schooling", unpublished Ph. D. dissertation, University of Chicago, 1965.
5. Hanoch, G., "An Economic Analysis of Earnings and Schooling", Journal of Human Resources, Summer 1967, pp. 310-329.
6. Denison, E.E. The Sources of Growth in the U.S. and the Alternatives Before U.S., op. cit.
7. Gounden, A.M.N., "Education and Economic Development", op. cit.
8. Gounden, A.M.N., Investment in Education in India, op. cit.
9. Hollister, R., "Education and Distribution of Income: Some Exploratory Forays", mimeographed draft, 1970.

III

THE NATURE OF INTERNAL RATES OF RETURN

In attempting to squeeze some meaning out of internal rates of return, it is important to remember that they capture at a moment of time the outcome of a whole series of past events and also reflect in part future events. We can think of these rates as reflecting the interaction between the supplies of and demands for different types of educated labour. The potential supply of labour was produced - and some is currently being produced - over several previous decades. The amounts produced as well as the facilities available to produce them reflect, in considerable part, the past "market" for educated manpower. Similarly, the configuration of potential demand for educated manpower has built into it past decisions about the educational intensity of production, reflecting the relative scarcity of available productive inputs, as well as those underlying factors that generate the final demand for output produced with educated manpower.

The supply and demand conditions also indicate in part future expectations. Even if present demand and supply appear to be in "balance", the prospects of sharp increases in either supply or demand in the near future are going to affect earnings levels and hence measured rates of return. Similarly, sharp expected increases in supply or demand can have effects on the costs of education and, hence, affect the rates of return independently of what might happen to future earnings levels.

Complicating all this is the fact that governmental policies may affect both the returns and the costs, and these policies may shift over time so that the prospects of disentagling the play of market forces from policy effects is difficult. In addition, other "imperfections" in the market will obviously have a bearing on the determinants of the rate of return through their effect on supply and demand.

The bare outlining of the factors which are "important" is of little help in interpreting rates of return. What it does suggest is the difficulty of the task. Until there is more research on the nature of the educational labour market, including its dynamics, we are not in a strong position to say much about the underlying determinants of internal rates of return or of changes in them.

USES OF INTERNAL RATES OF RETURN

Customarily, and notwithstanding our lack of knowledge, internal rates of return have been used to assess the payoff to educational investment - to the individual as well as to society - relative to the payoff yielded by other forms of investment. The objective has been to say whether added investment in education should or should not be made. Much less attention has been given to the fact that rates of return usually differ by levels of schooling and that the patterns of the rates have a bearing on the answer to questions regarding the profitability of schooling investments. Hence, the purpose here is to focus on the patterns of the rates of return, while giving little if any explicit attention to relative profitability.

PATTERNS OF RATES OF RETURN

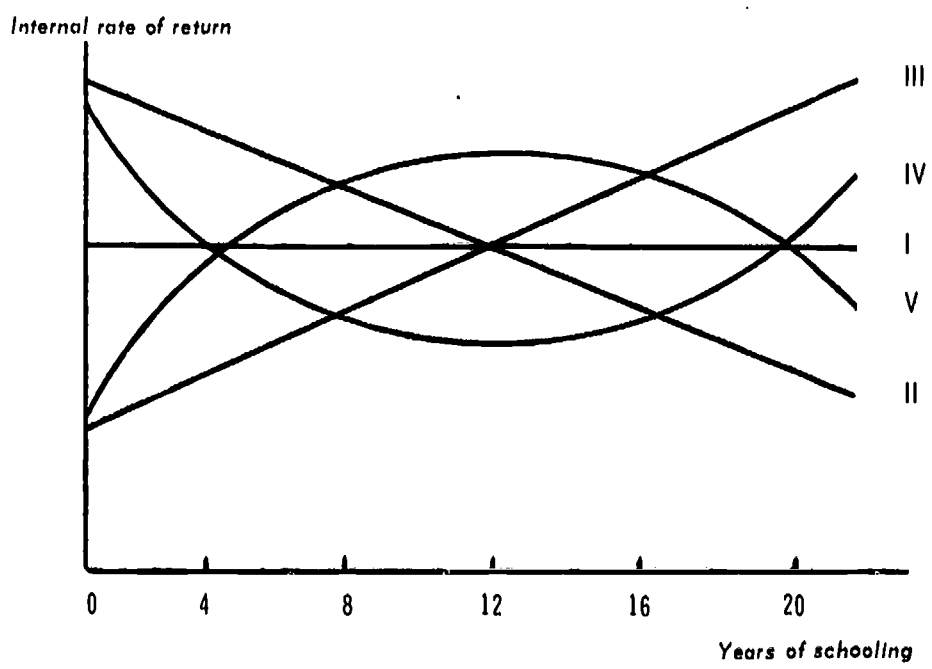
That any systematic patterns occur in the "incremental" rates of return across countries is not fully obvious from an inspection of the results of various countries; by "incremental" (sometimes referred to as "marginal") we refer to rates of return on each successive increment of schooling rather than on large blocks of schooling, e.g. from school entry through high school. It may be helpful, therefore, to set out several general types of patterns and then classify countries according to the type they fit most closely. The following five types of patterns appear capable of capturing most of the variations observed:

Type	Characteristic Pattern by Level of Educational Attainment
I	Constant rates of return across all levels of educational attainment.
II	Declining rates of return as level of educational attainment increases.
III	Rising rates of return as level of educational attainment increases.
IV	Declining and then rising rates of return.
V	Rising and then declining rates of return.

These patterns are shown in Figure 1.

Figure 1

ILLUSTRATIVE PATTERNS OF RATES OF RETURN TO EDUCATIONAL INVESTMENT



While greatly oversimplified, these patterns are of interest for their policy implications. Of course, what one concludes about the rates will depend on whether they refer to private or social rates of return[1]. Let us for purposes of discussion focus on the social rates of return. The first pattern (I) suggests a policy of indifference as to which level should be expanded or alternatively argues for an across-the-board change; the second (II) indicates need for a greater concentration at the early years of schooling; the third (III) suggests a concentration at the later years; and the fourth (IV) a concentration at both the early and later years; and the fifth (V) a concentration at the intermediate years. Actually, the policy implications of patterns III, IV, and V are more subtle than just indicated. Given that education is a sequential process, requiring the completion of lower levels of schooling prior to the higher levels, maximization of the rate of return to educational investment must take account not only of the incremental rates of return but also the rates of return over larger increments of education. To achieve a high rate of return at "say" the completion of high school, a steady flow of students through the lower and intermediate grades is essential even though the rates of return at these levels may be much lower. Too often, however, education is viewed as a series of discrete steps, without recognition that the higher level of attainment requires completion of a whole series of steps. This argues for a focus on both incremental and average rates of return.

Utilization of the typology set forth above produces the classification shown in Table 1, which is abstracted in turn from the results presented in Appendix Table B. This classification is based on "social" rates of return, i.e. on total resources invested; where social rates were not available it became necessary to use the private rates. The importance of the level of disaggregation on the classification scheme should be noted, the finer the breakdown by level of schooling, the greater is the possibility of undulating patterns. In some cases the rates for related levels of schooling (e.g. the first two years and the second two years of college) were averaged in determining the patterns. To assist in the classification, subtypes IVa and Va were established to take account of double reversals in the rate of return patterns.

It is striking to note that most of the results are of Type II and Type V, with their respective patterns of declining rates of return, and of rising and then declining rates of return. A smaller number of studies fall into Type IV, with only one country represented by Types I or III. Given the narrow range of schooling over which estimates are available for great Britain (II) (only the upper levels) and Greece (III) (only the lower levels), it is difficult to be certain about the appropriateness of the classifications of these countries. And the estimate for Venezuela (IV), one of the earliest, is probably subject to considerable error, given the assumptions made in calculating the rates of return. This leaves us with Types II, IV, and V. It is interesting to note that the Indian, Chilean, and Columbian studies fall into Types II and V. In view of the assumptions utilized in Harberger's Hyderabad study (V), we might want to attach somewhat less weight to it. It is more difficult, however, to explain away the dual classification of the Chilean and Columbian studies.

EXPLANATIONS OF RATES OF RETURN PATTERNS

The limited number of studies, their lack of comparability, and the approximate nature of the results preclude a systematic effort to explain the variety of patterns found and the placement of any country in a particular category. Nevertheless, we can offer some preliminary speculations. To help organize these speculations, let us advance several reasons why we would generally expect the rate of return patterns to be as they are[2].

1. Social rates of return are defined here as reflecting all monetary benefits and all costs - the direct and opportunity costs to students plus the other costs of education paid for by society.

2. Another intriguing question arises which cannot be discussed here: is it possible that questions regarding patterns and levels are intertwined, i.e. the relative pattern of the rates of return is related to the level of the rates of return to schooling or to the level of the rates offered by alternative investment opportunities ?

Table 1. CLASSIFICATION OF STUDIES ON RATES OF RETURN TO EDUCATIONAL INVESTMENT

TYPE	DESIGNATION OF STUDY
I	Constant rates of return
II	Declining rates of return
	Kenya, 1966 (Rogers)
	Bombay, 1956-67 (Kothari)
	India, 1960-61 (Gounden)
	India, 1960 (Selowsky)
	India Cities, 1964 (Reynolds)
	Great Britain, 1964 (Blaug)
	Santiago, Chile, 1962 (Bruton)
	Bogota, Columbia, 1963-66 (Selowsky)
	United States, 1949 (Hansen)
	United States, 1959 (Hanoch)
III	Rising rates of return
	Greece, 1960 (Liebenstein, in Bowles)
	Greece, 1964 (Liebenstein, in Bowles)
IV	Declining and then rising rates of return
	N. Rhodesia, 1960 (Baldwin)
	Israel, 1957-58 (Klinov-Malul)
	Venezuela, 1957 (Shoup)
	Canada, 1961 (Podoluk)
	IVa Declining, rising, and then declining rates of return
	Uganda, 1965 (Smyth and Bennett)
V	Rising and then declining rates of return
	Hyderabad, India, 1957 (Harberger)
	Imus, Cavite, Philippines, 1966 (Williamson and DeVoretz)
	Chile, 1958-59 (Harberger and Selowsky)
	Chile, 1964 (Selowsky)
	Bogota, Columbia, 1965 (Schultz)
	Columbia, 1961 (Gamacho, in Carnoy)
	Va Rising, declining and then rising rates of return
	Mexico, 1963 (Carnoy)
	Mexico, 1964 (Selowsky)

1) Influence of Cost-Return Relationships

The general tapering off of returns at the higher levels of schooling and the usually higher rates of return at the early levels of schooling suggest that literacy and all that goes with the completion of a few years of schooling pays off rather well, but that diminishing returns soon begin to set in. Why should

this be the case ? One important consideration is the cost structure. The total costs of education - in terms of both income foregone and the direct costs of schooling - rise rather steadily from elementary school up through college. Thus, on the basis of the overly simplified assumption that the returns to education rise by equal absolute amounts per year of schooling, while costs rise by equal percentage amounts, the rates of return will in general fall with more schooling. Of course, these underlying cost differences cannot provide a full explanation, since there is considerable variability in both the cost and return patterns.

2) Influence of Salary-Setting Mechanisms

The fact that rates of return do not always taper off more than they do may lie on the return side, particularly in developing countries. Because many highly educated people are in the employ of governments which ordinarily have rather rigidly prescribed salary schedules - schedules which are set by the educated portion of the population - the benefit stream may be abnormally high. Hence, this would hide the full extent of any decline. Offsetting this no doubt is the fact that government employment ordinarily brings with it a sizeable array of fringe benefits, most of which are not captured in money wage data. Moreover, the prestige that goes along with government employment may add further to any understatement of the "true" rate of return for the better-educated. Which of these forces is strongest, we simply do not know.

3) Supply and Demand Forces

The comments made thus far suggest that some constant, underlying forces are at work, and that these cut rather uniformly across countries and time. An alternative view is that the rate of return patterns reflect in part at least the impact of unique supply and demand forces, so that, if we had several sets of comparably calculated rates of return for different years, we would expect to find changing patterns. Put another way, the cross-section results may reflect disequilibrium conditions and so can be explained by reference to other events affecting supply, demand, or both. For example, changes in the rate of a country's growth may give rise to differential increases in the need for people by skill and amount of schooling. If growth accelerates, then the stock of highly educated workers, for example, may be insufficient, with the result that wage levels will be bid up. This will trigger a response, often a belated one, as additional people seek to obtain the types of education most needed; the result is to eventually push earnings and the rate of return back down again. Or, to take the opposite case, once a school system is geared up to a larger production level, and given inadequate information on the relative supply-demand situation, a larger number of people may enrol and eventually graduate than can be hired at prevailing salaries; the result will be either declining relative salaries, unemployment, or possibly both. In general, then, an increased demand for better-educated workers would tend to raise rates of return at the upper levels, and vice versa. This assumes that growth produces a rather education-specific pattern of demand for labour; such an assumption about the pattern of demand seems quite explicit in much of the work on educational planning in developing countries. Presumably, the validity of this assumption could be examined with the help of education-occupation and/or education-industry matrixes. An increased supply of educated people will, on the other hand, depress rates of return. What all this adds up to is the conclusion that rate of return patterns, given the way rates are usually calculated, are not unambiguous in the kind of information they provide.

IV

ASSOCIATION WITH EDUCATIONAL DISTRIBUTION DATA

Given the emphasis on educational distributions, it is appropriate to take the limited results we have and see how they "fit" with the distribution data. The data on educational distributions are of two types. One type represents "stocks" and the other "flows". Distributions of educational attainment of the population or work force fall into the stock category, whereas distributions of school enrolments and enrolment rates more closely represent "flows". We shall examine both of these types of data.

Any effort to relate educational distribution data to rate of return patterns is complicated by what we think rate of return patterns reflect. One view is that rate of return patterns reflect the outcome of past decisions, whereas another is that they provide a signal of what is likely to happen. Hence, we must quite carefully specify any expected relationships. If we adopt the former view, that the distributions strongly affect rate of return patterns, we would expect to find an inverse relationship between relative quantities (of educated people, students, etc.) and relative rates of return. On the other hand, the notion that the distributions reflect a response to current and expected conditions, signalled by rate of return patterns, would lead us to anticipate a positive association between relative quantities and relative rates of returns. We may find, therefore, that the rate of return patterns are consistent with one or the other of these two views of the role of rates of return.

Let us begin by looking at the distribution of educational attainment. We shall initially assume that rates of return reflect events of the recent past. This suggests that on average we can expect to find relative quantities and relative rates of return inversely related. More specifically, we would expect to find: 1) heavier concentrations of people with post-secondary educational attainment in countries of Types II and V, 2) heavier concentrations of people with elementary attainment in countries of Types III and V, and 3) heavier concentrations of people with secondary attainment in Type IV countries. The data in Table 2 are not consistent with 1) but they are generally consistent with 2) and 3). Thus, the distribution and rate of return data at least partially support the notion that rates of return reflect recent investment outcomes, as reflected by the distribution of relative quantities of educated manpower.

It should be clear that stock variables, such as data on the distribution of the educational attainment of the population or work force, are probably not entirely appropriate. Recent flows may at the margin have had a significant effect on earnings patterns and thus have altered rate of return patterns. Hence, flow variables are likely to be especially useful in casting light on rate of return patterns. If we take the distribution of students in school to represent the flow variable, and if we think of these flows as having a dominant effect on rate of return patterns, then we would expect to observe the same patterns 1), 2), and 3 discussed above. But Table 3 reveals clearly that the evidence does not support our expectations. Indeed, secondary education is most heavily concentrated in Type II countries, and elementary schooling shows a slightly heavier concentration in Type IV countries. On the other hand, there is no evidence of patterns the reverse of 1), 2), and 3) which would indicate people, through their enrolment patterns, are responding to rate of return patterns. We must therefore conclude that the flow data do not perform as we might have expected. Part of the difficulty may arise because the distribution of school enrolments is an imperfect measure of the flow of students out of the school system and into the labour

Table 2. PERCENTAGE DISTRIBUTION OF EMPLOYMENT BY LEVEL OF EDUCATIONAL ATTAINMENT

		PRIMARY	SECONDARY	POST-SECONDARY
I.				
II.	Kenya	-	-	-
	India	97	2	1
	Great Britain	52	40	7
	Columbia	-	-	-
	United States	53	27	19
III.	Greece	89	8	3
IV.	N. Rhodesia	-	-	-
	Israel	55	32	13
	Venezuela	-	-	-
	Canada	35	56	9
	IVa. Uganda	92	7	1
V.	India	97	2	1
	Philippines	80	14	6
	Chile	93	4	2
	Columbia	-	-	-
	Va. Mexico	93	5	2

SOURCE: Organisation for Economic Co-operation and Development, 1966.

Table 3. PERCENTAGE DISTRIBUTION OF SCHOOL ENROLMENT BY LEVEL OF SCHOOLING

		PRIMARY	SECONDARY	POST-SECONDARY
I.				
II.	Kenya	97	3	0
	India	62	36	2
	Great Britain	43	55	2
	Columbia	87	11	3
	United States	69	23	8
III.	Greece	77	21	2
IV.	N. Rhodesia	99	1	0
	Israel	85	12	3
	Venezuela	86	10	4
	Canada	77	20	3
	IVa. Uganda	78	21	1
V.	India			
	Philippines	78	13	8
	Chile	81	16	2
	Columbia	87	11	3
	Va. Mexico	97	6	2

SOURCE: UNESCO, Current School Enrolment Statistics.

force. Nor do such data tell much about either the possible queueing of people who desire either to enter the educational system or to proceed through ever-higher levels of the system. Nor do they tell about the effect of unemployment in whetting or dampening the desire on the part of young people to secure more education.

A further possibility is to look at enrolment rates rather than the distribution of enrolments, so as to better reflect the size of the flow relative to the stock of educated workers. The enrolment data are presented in Table 4. Our expectations would be identical with those earlier - high enrolment rates and low rates of return - if we expect the flows to have affected the current pattern of rates of return. The evidence in support of expectation 1) is not apparent, nor is it apparent in support of expectations 2) and 3).

Table 4. ENROLMENT RATES BY LEVEL OF SCHOOLING

		PRIMARY	SECONDARY	POST-SECONDARY
I.				
II.	Kenya	53	4	0.0
	India	32	22	1.4
	Great Britain	61	105	4.9
	Columbia	46	17	1.8
	United States	81	80	33.9
III.	Greece	62	44	3.5
IV.	N. Rhodesia	-	-	-
	Israel	78	49	10.2
	Venezuela	71	30	5.1
	Canada	90	61	9.0
	IVa. Uganda	32	6	0.2
V.	India	32	22	1.4
	Philippines	56	25	9.7
	Chile	69	35	3.5
	Columbia	46	17	1.8
	Va. Mexico	57	11	3.1

SOURCE: Columns 1-2, UNESCO, 1966, Table 4, and Column 3 based on data from UN, Demographic Yearbook; UNESCO, 1966.

V

CONCLUSION

This analysis has brought out the fact that the patterns of rates of return to educational investment for the most part either fall with more schooling or first rise and then fall. But we have been able to show at best a weak association between these patterns and the educational structure, either as reflected in educational attainment data or in enrolment patterns. Several factors account for these rather disappointing findings. First, the models that we possess to explain the observed distributions of most variables are extremely primitive; although we can fit functions of one kind or another to such distributions, we are unable to say much about the forces which generate them. Second, in the absence of any substantial analytical scheme, one can only search for empirical regularities, as has been tried here. But because the sample of rate of return studies is still so small, this search has necessarily been a crude and exploratory one.

In future work on this topic, several steps must be taken. The first is to augment the supply of rate of return studies, preferably by assembling additional studies which have already been completed but are not included here. The second and more difficult is to produce more and better data on educational distributions so as to fill the existing information gaps. We need additional data on both stocks and flows. In particular, we require data that reflect what is going on at the key junctures in the educational system (continuation rates by level of schooling), and what is happening at the point where the educational system and the labour market join together (unemployment rates by level of school attainment for new and recent entrants into the labour force). Finally, we need unemployment rates by level of school attainment for those people already in the labour force. The third and still more difficult task is to develop a comprehensive set of hypotheses that seek to explain differences in rate of return patterns so that these hypotheses can be subjected to empirical testing. This requires going beyond using only educational distribution data. In addition, variables reflecting demand conditions must be introduced, so as to capture the critical supply and demand factors which are at work. As a result of such work it should be possible for us to gain a better understanding of the determinants of rate of return patterns and their link to the underlying quantities of educated manpower. In the meantime, we are left with an intriguing set of observations that begs for an explanation.

Appendix Tables A and B

PATTERNS OF RATES OF RETURN TO INVESTMENT IN EDUCATION: SOME INTERNATIONAL COMPARISONS

W. Lee Hansen
University of Wisconsin

March 1970

Table A. SUMMARY OF STUDIES ON RATES OF RETURN TO EDUCATIONAL INVESTMENT

AUTHOR, DATE, AND REFERENCE	YEAR OF STUDY, LOCATION; AND TYPE OF WORKER	TYPE OF DATA	METHODS USED IN CONSTRUCTION OF BENEFIT STREAMS	CONTROLS AND ADJUSTMENTS, AND ADJUSTMENTS FOR ECONOMIC GROWTH
		AFRICA		
Baldwin, (3)	1960, Northern Rhodesia	Urban males	Average wage levels without regard to age.	
Rogers, 1968 (39)	1966, Kenya; civil servants and teachers.	Government pay scales for civil servants and teachers.	No information.	Assumed a standard income by educational level.
Smyth and Bennett, 1966 (48)	1965, Uganda	Public salary scales grossed upward to account for higher salary scales in private sector.	Age-earning curves are linear exponential curves derived by inspection from the Government scales.	Adjustment for alternative entry to wage or farm employment.
		ASIA		
Gounden, 1965 (16, 17)	1960-61, India, urban males and engineers.	NCAER survey of 5,000 urban males, CSIR study of 4,000 engineers.	Age-income streams taken directly from average annual income tabulated by age and education level.	Arbitrary adjustment of 50% to account for non-educational determinants of income.
Harberger, 1963 (22)	1957, Hyderabad, India; male earners.	Sample of 5,885 earners distributed by earnings and schooling.	Assumed age-earning profiles with the following form: 1) peak in earnings reached at a later age for successive educational levels 2) peak earnings were a higher fraction of average earnings for successive educational levels.	Adjusted for labour force participation.
Kothari, 1966 (31)	1956-57, Bombay City, India.	3% random survey of tenaments in Greater Bombay, earnings by education level. (not by age).	Estimated an age-earning profile with similar assumptions as Harberger plus assumption that earnings decreased after age 55.	Correction for survival rates of cohorts by educational level would reduce rate of return by about 2%.
Reynolds, 1968 (38)	1964, Bombay, Jamshedpur, Madras, and Rowkela; urban males.	Sample of 1,800 production workers in steel and metal-working firms.	Age-education-earnings data taken directly from survey results.	
Selowsky, 1967, (45)	1960-61, India; urban males and engineers.	Gounden's data.	Gounden's age-income profiles (including costs).	Attributed 100% of the wage differential to education rather than one half as did Gounden.

Table A. (cont'd) SUMMARY OF STUDIES ON RATES OF RETURN TO EDUCATIONAL INVESTMENT

AUTHOR, DATE, AND REFERENCE	YEAR OF STUDY, LOCATION; AND TYPE OF WORKER	TYPE OF DATA	METHODS USED IN CONSTRUCTION OF BENEFIT STREAMS	CONTROLS AND ADJUSTMENTS, AND ADJUSTMENTS FOR ECONOMIC GROWTH
		ASIA (cont'd)		
Williamson and DeVoretz, 1967, (54)	1966, Imus, Cavite (Philippines); male heads of households.	A Population Institute survey. The sample size was 1,063 and gives average annual income by age and level of schooling.	Regressions made of income on the age variable for each education level to yield an age-earnings profile.	Attributed 50% of differential to education. Used mortality adjustment.
		EUROPE AND MIDDLE EAST		
Blaug, 1967, (7,8)	1964 and 1967, Great Britain a) male workers; b) professional, managerial and skilled workers.	2 surveys: a) 1964, random sample of 6,500 male heads of households; b) 1967, sample of 2,800 workers in British auto industry and 4 large electrical engineering firms.	1st survey: earning differentials by present age are calculated; 2nd survey: age-earnings profiles are obtained from the data.	Assumed that 0.6 of observed differential associated with extra education is due to education for 1st survey. No adjustment for ability, social class, etc. used for 2nd survey.
Klinov-Malul, 1966 (30)	1957-58, Israel, a) Jewish urban workers - heads of households b) professionals	a) survey of 3,000 families - wage and salary income by education. b) sample survey of 4 professionals (1,000).	Calculation of the net effect of education for the population of each continent-residence combination; sometimes utilizing standardization and least squares methods.	Results standardized for continent of origin and length of residence. Alternate calculations of individual returns at a 3% annual growth in GNP.
Leibenstein, 1967, (32)	1960, 1964, Athens, Greece; male workers and female workers.	Sample size of 2,700, plus salary information on public workers and professional organizations.	Age-earnings profiles are constructed. Taken directly from data.	Assumes alternative annual growth rates of 0, 4, and 5%.
		LATIN AMERICA		
Bruton, 1967, (9)	1962, Greater Santiago, Chile; male members of labour force.	Data on education, age, and wage income (2,500 observations - a number of which were excluded).	Earnings are regressed on education to yield a wage-education relationship. Age-earnings profiles are developed.	
Camacho, 1964, (11)	1961, Columbia urban, males.	No information.	No information.	
Harberger and Selowsky, 1966 (23)	1958-59, Chile, male and female.	Sample size is not specified.	Age-earnings profiles are not constructed, rather they assume that the differential between the earnings of the education groups is relatively constant as a function of age.	

Table A. (cont'd) SUMMARY OF STUDIES ON RATES OF RETURN TO EDUCATIONAL INVESTMENT

AUTHOR, DATE, AND REFERENCE	YEAR OF STUDY, LOCATION; AND TYPE OF WORKER	TYPE OF DATA	METHODS USED IN CONSTRUCTION OF BENEFIT STREAMS	CONTROLS AND ADJUSTMENTS, AND ADJUSTMENTS FOR ECONOMIC GROWTH
		LATIN AMERICA (cont'd)		
Schultz, 1968, (41)	1965, Bogota, Columbia	Sample Size: 684 men, 314 women. Both hourly and weekly earnings by age and education.	Regression estimates of age-log earnings profiles by level.	Adjustment for migration to Bogota.
Shoup, 1959, (47)	1950's, Venezuela	Sample data	Based on assumed shapes of age-earnings profiles.	
Selowsky, 1967, (45)	1964, Chile; urban males.	No information.	Age-earning profile in "Encuesta Nivel di Vida" centro de Planificacio'n Economica, Universidad di Chile, 1964.	
Selowsky, 1968, (46)	1963-66, Bogota, Columbia; urban males and females.	Hourly wage data by schooling and age taken from unemployment samples. 10,715 observations.	Age-earnings profiles from data.	Five versions calculated with adjustments for L. F. participation rates, unemployment, changes over time. Version 5 - adjustment for growth in L. F. and gross domestic product.
		NORTH AMERICA		
Carnoy, 1964, (12, 13)	1963, Mexico City, Puebla, and Monterey, Mexico; male wage earners in 8 occupational classes.	3,901 observations, data on wage and salary, schooling in years, age, father's occupation, discipline of study, and industry.	Earnings are regressed on schooling, age, occupations, father's occupation, etc. ; sample then divided into schooling levels which permitted analysis within each category. From these results lifetime earnings streams are constructed.	Controlled for father's occupation plus other variables (industry, city).
Hanoch, 1967, (21)	1959, United States, males, white and nonwhite for North and South.	1960, Census data.	Used mean age-education-income data, adjusted for variables correlated with age.	
Hansen, 1963, (18)	1949, United States, males.	1950, Census data.	Used mean age-education-income figures.	Adjusted for mortality.

Table A. (cont'd) SUMMARY OF STUDIES ON RATES OF RETURN TO EDUCATIONAL INVESTMENT

AUTHOR, DATE, AND REFERENCE	YEAR OF STUDY, LOCATION; AND TYPE OF WORKER	TYPE OF DATA	METHODS USED IN CONSTRUCTION OF BENEFIT STREAMS	CONTROLS AND ADJUSTMENTS AND ADJUSTMENTS FOR ECONOMIC GROWTH
		NORTH AMERICA		
Podoluk, 1965, (36)	1961, Canadian males.	1961, Census data.	Used average age-education-earnings data.	
Selowsky, 1967, (45)	1963-64, Mexico.	Used Carnoy's sample and a sample from the "Direccion de Muestro".	Adjustment for expected labour force participation, unemployment, and survival rate.	Adjusted for expected annual growth rate of wages of level of schooling.

Table B. SUMMARY OF RATE OF RETURN ESTIMATES

CITY, COUNTRY, AND YEAR	LEVEL OF SCHOOLING	RATE OF RETURN		SPECIAL NOTES
		PRIVATE	SOCIAL	
	AFRICA			
Kenya, 1966 (39) (Rogers)	4th form plus	58		Assumption of no unemployment; earnings include housing subsidy; persons begin entering the work force at age 18 and retire at age 55; private direct costs for higher levels in Kenya is zero; costs include foregone earnings and the individual and the state direct costs.
	9 months governmental training/4th form			
	6th form/4th form	negative		
	4th form plus 2 years primary teacher training/ 4th form	10.6		
	4th form plus 3 years secondary teacher training/ 4th form plus primary teacher training	39.0		
	4th form plus secondary teacher training/6th form	42.0		
	university/4th form plus secondary teacher training	19.0		
	university/6th form	26.0		
	university/4th form plus primary teacher training	24.5		
Northern Rhodesia, 1960, (3) (Baldwin)	Standard I/0		13	Assumes constant earnings differences; costs include foregone income and school costs; discounts over 45 years, in recognition of upward bias imported by assumption of constant earnings differences.
	Standard II/I		11	
	Standard III/II		4	
	Standard IV/III		15	
	Standard V/IV		16	
	Standard VI/V		22	
	Estimates from present value results			
Uganda, 1965 (48) (Smyth and Bennett)	primary (7)/(0)		66	Earning flows for primary educated manpower are only estimates; cost data include foregone earnings, capital repayment on education plant, recurring operating costs, interest foregone (and the cost of educating those who do not pass their examinations).
	CSC(11)/P(7)		22	
	HSC(13)/CSC(11)		78	
	university(16)/HSC(13)		12	
	ASIA			
Imus, Cavite; Philippines, 1966 (54) (Williamson and DeVoretz)	elementary(7)/illiterates	9	8	Cost data is from 1965 extrapoled to 1966; it includes direct expenditures and earnings foregone for the private estimates plus government expenditures for social costs.
	high school/elementary	29	21	
	college/high school	12	11	
Bombay City, India, 1956-57, (31) (Kothari)	high school (12)/middle (8)	-	20	It was necessary to estimate the age structure of the earners and to isolate the influence of business and commercial owners under the assumption that education only marginally influenced their earnings.
	college (17)/(12)	14	13	
	engineering (17)/(12)	25	22	
	arts and sciences (17)/(12)		10	

Table B. (cont'd) SUMMARY OF RATE OF RETURN ESTIMATES

CITY, COUNTRY, AND YEAR	LEVEL OF SCHOOLING	RATE OF RETURN		SPECIAL NOTES
		PRIVATE	SOCIAL	
	ASIA (cont'd)			
Hyderabad, India, 1957 (22) (Harberger)	secondary (12)/primary (8) college and university (18)/S (12) secondary plus college/P (8)		11.9 16.9 15.0	Sample was heavily weighted with younger people so sample was reweighted: data referred to people with some primary, etc. so it was necessary to estimated average income of completers; assumptions made were likely to produce overestimates for rates of return; cost data: earnings foregone and assumed direct costs (conservative).
India, 1960 (45) (Selowsky)	(5)/(2) (8)/(5) (11)/(8) (15)/(11) (17)/(15)	23.5 17.7 16.4 11.6 14.7	21.2 19.9 18.9 16.2 16.0	
India, 1960-61 (16,17) (Gounden)	literates/illiterates primary (5)/literate middle (8)/P (5) matriculates (12)/M (8) bach. degree (15)/M (12) engineering (17)/M (12) engin. (17)/bach. (15)	30 23.0 13.0 10.0 8.1 13.5 20.5	15.9 17.0 11.8 10.3 7.0 9.8 9.7	It was not always possible to isolate earnings figures from income figures; those with primary education and below enter the labour force at age 12 and retire at age 60; assumes full employment; cost data include direct expenditures, depreciation of physical assets, imputed value of interests, and foregone earnings; gross investment in education forms 8.5% of adjusted NY and 44.1% of gross physical capital formation.
India Cities, 1964 (38) (Reynolds)	primary (5)/literate middle (8)/primary (5) matriculate (12)/middle (8) two years college (24)/ matriculate (12)	21.0 12.0 11.4 4.4	14.5 9.1 7.0 1.8	Cost data based on updating of Gounden data from 1961; based on earnings data to age 60.
	EUROPE			
Great Britain, (6,8) 1964 (Blaug)	terminal education age (15-18) (15-21)	13 14	12.5 8	1st survey: no distinction between types of schooling or between full- and part-time schooling; breakdown by age is too large to allow standardization.

Table B. (cont'd) SUMMARY OF RATE OF RETURN ESTIMATES

CITY, COUNTRY, AND YEAR	LEVEL OF SCHOOLING	RATE OF RETURN		SPECIAL NOTES
		PRIVATE	SOCIAL	
Great Britain, (cont'd) 1967	Level of Qualification Royal Society of Arts, and City and Guild Institute of London preliminary certificates	5 (5)	5.5 (5.5)	2nd survey: more detailed information on type and level of education; social rates of return are calculated from before-tax earnings with cost data as the total resource costs including income foregone; the private rates of return are calculated from after the earnings and reflect only private costs.
	General Certificate - ordinary level, RSA advanced and CGL intermediate certificate	8 (/5)	8.5 (15)	
	General Certificate - advanced level, ordinary national certificate and ordinary national diploma	8 (1.5)	7 (0)	Figures not in parentheses refer to comparison with school-leaving age. Figures in parentheses refer to comparison with previous level.
	Higher National Certificate, CGL full technical certificate.	9.5 (15)	7.5 (12)	
	university degree, higher national diploma	8.5 (7)	6 (5)	
	university degree (honours), diploma in technology	9.5 (10)	8 (8)	
Greece (32) 1960	12/6 (male) 15/6 " 15/12 " 15/6 (female)		4.5 (9) 6 (10) 8 (12.5) 3 (7)	The survey collected data on each worker's age, years of education (both technical and general), monthly earnings, and occupation.
1964 (Liebenstein)	12/6 (male) 15/6 " 15/12 " 12/6 (female)		3 (7) 5 (9.5) 8 (12.5) 5 (9.5)	Figures in parentheses assume a 4% rate of economic growth.
Israel, 1957-58 (30) (Klinov-Malul)	(Present Values, IL thousands; 8% discount rate)			Income tax is used as a partial measure of returns to society; the income of younger professional workers is rising relative to older workers. This change in the structure of income by age is somewhat peculiar to Israel (due to immigration) and is reflected even more strongly in the present value of incomes; cost data include expenditures by society (salaries and wages, books

	PRIVATE	SOCIAL	SOCIAL ADJUSTED FOR 3% GNP GROWTH RATE
Primary	11.3	9.4	
Secondary		-0.8	+5.6
(discount rate 10%)	(-4.0)	(-4.0)	
Higher Education		6.1	2.7
Engineers	1.3	5.3	16.4
Lawyers	-1.0	-0.2	3.1
CPA's	42.9	25.9	82.7
Physicians	-24.9	.6	-13.6

Table B. (cont'd) SUMMARY OF RATE OF RETURN ESTIMATES

CITY, COUNTRY, AND YEAR	LEVEL OF SCHOOLING	RATE OF RETURN		SPECIAL NOTES
		PRIVATE	SOCIAL	
Israel, 1957-58 (cont'd)				and materials, depreciation in buildings) and the private expenditures of individuals (income foregone, tuition and fees, and books and materials).
	LATIN AMERICA			
Chile, 1958-59 (23) (Harberger and Selowsky)	primary (average 5.5 years)/ none "special" secondary (average 8.5 years) secondary (average 11.5 years) university		24.0 29.0 16.9 12.2	Cost data is from $ and # Yver, "The Cost of Education in Chile". (Universidad Catolica de Chile, 1959). (mimeographed).
Chile, 1964 (45) (Selowsky)	2/0 7.7 4/2 19.1 6/4 24.8 8/6 12.4 12/10 22.9		7.7 13.4 17.2 16.0 15.3	Used cost data of Yver.
Santiago, Chile, 1962 (9) (Bruton)	primary (6)/none secondary (12)/primary university (17)/secondary (figures in parentheses include the costs of educating all persons attending school whether they finish or not)	18 18 14	(16.5) (18) (14)	Income data is a byproduct of a survey on unemployment conducted by the Instituto de Economia, Universidad de Chile; cost data is from Raul E. Yver's study and include direct current outlays in teacher salaries, books and supplies plus rental value of school building, grounds and equipment plus foregone earnings.
Bogota, Columbia (46) 1963-66. (Selowsky)	(all RofR are social) primary (3)/illiterates primary (5)/illiterates bachillerato (11)/B (8) B (11)/B (10) university (16)/B (11) university (14)/B (11)	2 32 33 23 21 6 neg.	3 26 28 24 21 6 neg.	5 28 30 25 23 7 neg.	Costs include earnings foregone, payments to teachers and depreciation and interest of educational equipment. Version: 2) adjusted for L.F. participation rates; assume full employment 3) adjusted for unemployment which is substantial in low wage groups 5) assumption of growth over time in labour force and GDP.
Bogota, Columbia, 1965 (41) (Schultz)	primary (5)/none secondary (11)/P (5) vocational (8)/P (5) university (16)/S (11)	18.2 34.4 51.6 4.6	15.3 26.5 35.4 2.9	

Table B. (cont'd) SUMMARY OF RATE OF RETURN ESTIMATES

CITY, COUNTRY, AND YEAR	LEVEL OF SCHOOLING	RATE OF RETURN		SPECIAL NOTES
		PRIVATE	SOCIAL	
Columbia, 1961 (11) (Camacho)	primary (5) technical secondary (11) general secondary (11) university	20 19 30 19		
Venezuela, 1957 (47) (Shoup)	primary (6)/illiterates 7-11 12-15		82 17 23	Earnings foregone are not included in cost estimates of the primary rate which would lower the rate to approx. 30%.
	NORTH AMERICA			
Canada, 1961 (36) (Podoluk)	elementary secondary university	16.3 16.3 19.7		Assumes no unemployment.
Mexico, 1963 (12,13) (Carnoy) unadjusted	2-4 5-6 7-8 9-11 12-13 14-16	21.1 48.6 36.5 17.4 15.8 36.7	17.3 37.5 23.4 14.2 12.4 29.5	Private costs include direct expenditures on tuition, books, transportation, supplies, etc. plus earnings foregone; social costs originate in a study of public expenditure on formal schooling in Mexico - 1940-1962; it includes implied rent and depreciation changes for buildings.
father's income constant	2-4 5-6 7-8 9-11 12-13 14-16	15.2 44.9 31.0 15.2 14.6 39.5	12.8 34.5 20.6 12.3 11.4 31.5	
Mexico, 1964 (46) (Selowsky)	4/0 (marginal) 6/5 7-8/6 9-11/7-8 12-13/9-11 14-16/12-13	17.8 37.3 24.0 15.1 14.4 29.9	17.3 24.3 22.5 21.4	Adapted from Carnoy's results. (total - Carnoy's data)
United States, 1949 (18) (Hansen)	2/0 6/2 8/6 10/8 12/10 14/12 16/14	12.7 18.6 6.2 18.7	89 14.5 29.2 9.5 13.7 5.4 15.6	Assumes no unemployment. Costs include all of usual components. 11.4 10.2
United States, 1959 (20,21) (Hanoch)	4/0 (white North) 6/4 8/6 10/8 12/10 14/12 16/14	100.0 21.8 16.3 16.0 7.1 12.2 7.0	89.0 6.0 10.0 12.0 7.0 7.0 5.0	Assumes no unemployment; no adjustment for mortality. (non-white South)

REFERENCES

1. Alchian, A. A., "The Rate of Interest, Fisher's Rate of Return Over Cost, And Keynes Internal Rate of Return", American Economic Review, December 1955, pp. 938-943.

2. Bailey, M. J., "Formal Criteria for Investment Decisions", Journal of Political Economy, October 1959, pp. 476-483.

3. Baldwin, R. E., Economic Development and Export Growth, Berkeley: University of California Press, 1966.

4. Becker, G. S., "Underinvestment in College Education", American Economic Review, May 1960, pp. 346-354.

5. Becker, G. S., Human Capital, Princeton: Princeton University Press, 1964.

6. Blaug, M., "The Rate of Return on Investment in Education in Great Britain", Manchester School, September 1965, pp. 205-262.

7. Blaug, M., "Approaches to Educational Planning", Economic Journal, June 1967, pp. 262-288.

8. Blaug, M., "The Private and the Social Returns on Investment in Education: Some Results for Great Britain", The Journal of Human Resources, Summer 1967, pp. 330-346.

9. Bruton, H. J., "The Productivity of Education in Chile", Center for Development of Economics, Williams College, Research Memorandum No. 12, 1967.

10. Bowles, S., Planning Educational Systems for Economic Growth, Harvard: Harvard University Press, 1969.

11. Camacho, G. F., "Rendimento de la Inversion en Educacion en Colombia", Bogota: CEDE, Universidad de los Andes, July 1964; Cited in Carnoy (1967) p. 363.

12. Carnoy, M., The Cost and Return to Schooling in Mexico: A Case Study, unpublished Ph. D. dissertation, University of Chicago, 1964.

13. Carnoy, M., "Rates of Return to Schooling in Latin America", Journal of Human Resources, Summer 1967, pp. 359-374.

14. Denison, E. E., The Sources of Growth in the U. S. and the Alternatives Before Us, Washington, D. C.: Committee for Economic Development, 1962.

15. Fabricant, S., Basic Facts on Productivity Change, New York: National Bureau of Economic Research, 1959.

16. Gounden, A. M. N., Education and Economic Development, unpublished Ph. D. dissertation, Karukshetra University, India, 1965.

17. Gounden, A. M. N., "Investment in Education in India", Journal of Human Resources, Summer 1967, pp. 347-358.

18. Hansen, W. L., "Total and Private Rates of Return to Investment in Schooling", Journal of Political Economy, April 1963, pp. 128-41.

19. Hansen, W. L., Comment on Harberger's Estimate of Rate of Return to Investment in Education in India, mimeographed, August 7, 1963.

20. Hanoch, G., Personal Earnings and Investment in Schooling, unpublished Ph. D. dissertation, University of Chicago, 1965.

21. Hanoch, G., "An Economic Analysis of Earnings and Schooling", Journal of Human Resources, Summer 1967, pp. 310-329.

22. Harberger, A. C., "Investment in Man Versus Investment in Machine", in C. A. Anderson and M. J. Bowman, eds., Education and Economic Development, Chicago: Aldine, 1965, pp. 11-50.

23. Harberger, A. C., and Selowsky, M., The Key Factors in the Economic Growth of Chile: An Analysis of the Sources of Past Growth and of Prospects for 1965-70, mimeographed, Cornell University, 1966. Also cited in Carnoy, 1967, p. 368.

24. Harbison, F., and Myers, C. A., Education, Manpower, and Economic Growth, New York: McGraw-Hill, 1964.

25. Hirschleifer, J. A., "On the Theory of Optimal Investment Decision", Journal of Political Economy, August 1958, pp. 329-352.

26. Hollister, R., Education and Distribution of Income: Some Exploratory Forays, mimeographed draft, 1970.

27. Jean, W. H., "Reply", Journal of Business, March 1969, pp. 99-100.

28. Johnson, H. G., "Observations" in Organisation for Economic Co-operation and Development, The Residual Factor in Economic Growth, Paris 1964, pp. 219-227.

29. Kendrick, J. W., Productivity Change in the U. S., Princeton: Princeton University Press, 1961.

30. Klinov-Malul, R., The Profitability of Investment in Education in Israel, Jerusalem: Falk Foundation, 1966.

31. Kothari, V. N., Returns to Education in India, mimeographed, University of Baroda, 1966.

32. Leibenstein, N., "Rates of Return to Education in Greece", Development Advisory Service, Harvard University, Economic Development Report No. 94, September 1967, p. 13. Also cited in Bowles, 1969, pp. 153-156.

33. Myrdal, G., Asian Drama, New York, 1968, Vol. III.

34. Organisation for Economic Co-operation and Development, Statistics of the Occupational and Educational Structure of the Labour Force in 53 countries, Paris 1969.

35. Pandit, H. N., Productivity of Investment in Education: the Conceptual Framework, Techniques, and Major Findings, New Delhi: National Institute of Education, 1969.

36. Podoluk, J. R., Earnings and Education, Ottawa: Dominion Bureau of Statistics, 1964.

37. Psacharopoulos, G., "Estimating Shadow Rates of Return to Investment in Education", Journal of Human Resources, Winter 1970, pp. 34-50.

38. Reynolds, M., "Some Evidence on the Return to Schooling in Indian Industry", unpublished manuscript, October 1968.

39. Rogers, D. C., The Returns to Investment in Higher Levels of Education in Kenya, Institute for Development Studies, University College, Nairobi, 1968.

40. Rogers, D. C., Private Rates of Return to Education in the United States: A Case Study, unpublished Ph. D. dissertation, Yale University, 1967.

41. Schultz, T. P., "Returns to Education in Bogota, Columbia", RAND Corporation, Memorandum RM-5645-RC/AID, September 1968.

42. Schultz, T. W., "Investment in Man: An Economist's View", Social Science Review, June 1959, pp. 108-117.

43. Schultz, T. W., "Education and Economic Growth", In N. B. Henry ed., Social Forces Influencing American Education, Chicago: University of Chicago Press, 1961, Part II, pp. 46-88.

44. Schultz, T. W., "The Rate of Return in Allocating Investment Resources to Education", Journal of Human Resources, Summer 1967, pp. 293-310.

45. Selowsky, M., "Education and Economic Growth: Some International Comparisons", Project for Quantitative Research in Economic Development, Harvard University, Economic Development Report No. 83, November 1967.

46. Selowsky, M., "The Effect on Unemployment and Growth on the Rate of Return to Education: The Case of Columbia", Project for Quantitative Research in Economic Development, Harvard University, Economic Development Report No. 116, 1968.

47. Shoup, C. S., The Fiscal System of Venezuela, Baltimore: Johns Hopkins Press, 1959.

48. Smyth, J. A., and Bennett, N. L., "Rates of Return on Investment in Education: A Tool for Short-Term Educational Planning, Illustrated with Uganda Data", In G. Z. F. Bereday and J. A. Lauwerys, eds., 1967 World Year Book on Education, "Education Planning", New York: Harcourt, Brace and World, 1967, pp. 299-322.

49. Solow, R. M., Capital Theory and the Rate of Return, Amsterdam: Rand McNally, 1963.

50. UN, Demographic Yearbook, (various years).

51. UNESCO, World Survey of Education, Paris 1966, Vol. IV.

52. UNESCO, Current School Enrolment Statistics, (various years).

53. Williamson, J. G., "Economic Growth in the Philippines: 1947-1965: The Role of Traditional Inputs, Education, and Technical Change", Institute of Economic Development and Research, School of Economics, University of the Philippines, Discussion Paper No. 67-8, September 12, 1967.

54. Williamson, J. G., and DeVoretz, D. J., "Education as an Asset in the Philippine Economy", In M. B. Conception, ed., Philippine Population in the Seventies, Manila: Community Publishers, Inc., 1969, pp. 133-168.

Paper 5

THE SEARCH FOR EQUITY IN THE PROVISION AND FINANCE OF HIGHER EDUCATION[1]

by

W. Lee Hansen and Burton A. Weisbrod

1. Based on material in chapters 4 and 6 in W. Lee Hansen and Burton A. Weisbrod, Benefits, Costs and Finance of Higher Education, Chicago, Markham Publishing Co., 1969. See also W. Lee Hansen and Burton A. Weisbrod, "The Distribution of Costs and Direct Benefits of Public Higher Education: The Case of California", Journal of Human Resources, Spring 1969, pp. 176-191.

INTRODUCTION

Who should be eligible for public higher education? Should those young people who are not eligible - or, if eligible, are unable or unwilling to go to college - be deprived of the public subsidies obtained by the college goers? What can be said about the actual distribution of public subsidies for higher education - that is, who actually receives them? And who pays for them? These and related questions are explored in this paper, which is addressed to the subject of equity in the provision and financing of higher education. First, some conceptual issues are treated, and then a newly-available body of data is analysed with the objective of determining how the benefits and costs of public higher education are actually shared in our most populous state, California.

I

EQUITY AND EFFICIENCY: CONCEPTUAL ISSUES

How should public higher education be financed? There are actually two separatable questions: one, who should bear the costs of public higher education; two, how the portion of costs that is borne by students should be paid. More precisely, the question of who should pay involves determining the share of costs to be paid by students versus taxpayers[1]. The question of how students should pay relates directly to the tuition issue, but "the tuition" is not a simple concept. Should tuition be the same for all students? Whatever the level or levels of tuition, should it be paid at the time the education is received, or later? Should the level of tuition be determined at the time the education is received, or should the amount be contingent on future benefits?

The nature of these choices will be described more fully below. While our primary concern in this paper is with equity, we recognize that sound public policy should also strive for efficiency in the use of resources. Thus, we begin with a discussion of what we mean by "efficient" and "equitable" solutions to educational finance questions.

By "economic efficiency" we mean the degree of success of higher education in producing outputs (trained students, for example) that are more valuable than the resources used up in the process of production. In the economy at large, the value of output is generally measured by what people are willing to pay for it, and, so, as a first approximation, the value of college education may be measured by the increased salaries that employers are willing to pay for workers who are college educated rather than only high school educated. Efficiency, in the present context, can thus be thought of in terms of the amount by which national income (or Gross National Product) is raised by higher education[2].

But, which people receive this increased income? And who pays the costs of the resources - teachers, classrooms, laboratories, etc. - that are required to produce the increased income? In other words, how "fairly" are the additional income and the costs of public higher education shared? This is the issue of equity.

The distinction between efficiency and equity is essential if we are to come to grips intelligently with difficult issues of public policy. Higher education may be found to be efficient in raising incomes, but the method of financing higher education might be regarded as inequitable. By contrast, it may be felt that higher education is being financed equitably, but that it is really not an efficient way to use resources - there being better ways to increase people's real incomes (such as by devoting more resources to improving technology). Of course, there are intermediate positions, in which various degrees of inefficiency and inequity are adjudged to exist.

1. There is, of course, the question of how the taxpayers portion of the costs is to be shared among various groups, but we do not deal with this matter.

2. But this is only a first approximation of the full value of college education, because it disregards benefits that arise in ways other than through the job market. We return to this point below. On the other hand, it disregards the fact that the increased salaries reflect not only the effects of schooling but also of the generally greater ability of those students who have opted for more schooling.

Debate over issues in higher education finance can only be fruitful if there is a recognition of when, and to what extent, the dispute centres on factual matters of efficiency, and when it centres on value judgments regarding the fairness of the distribution of benefits and costs. This is not to deny, however, that both classes of issues are difficult to resolve, for the factual data relevant to assessment of efficiency are difficult to find, as is consensus on what should be regarded as equitable.

The social objectives of efficiency and equity are in fact quite likely to conflict, thereby complicating the issue. Consideration of efficiency might suggest that higher education should be provided to some young people but not to all; implicit is the widely held assumption that not everyone can benefit significantly from higher education. But there is still the equity question: it is "fair" for some youngsters to receive public subsidies while others do not? An efficient allocation of resources can be inequitable.

And, an equitable allocation of resources can be inefficient. If, for example, every youngster were not only offered the opportunity to go to college for four years, but were required to go, then all college-age people would receive a similar public subsidy. But if this is more equitable it is doubtless less efficient, for not everyone is likely to benefit enough to cover the costs of resources required to educate them. The conflict between equity and economic efficiency in higher education planning appears to be a genuine one; we do not attempt to resolve it here, but rather focus on the issue of equity itself.

Throughout this paper our attention is directed primarily to decision-making in the public sector. This orientation is somewhat artificial. The fact that there exists a private as well as a public sector in higher education means that success in devising an efficient and equitable finance system for public higher education does not assure either efficiency or equity for the higher education system as a whole, although our analysis of efficiency and equity issues in the financing of public higher education is applicable to private as well as public higher education. The question of what separate and distinct roles ought to be fulfilled by the public and private sectors in higher education is an important one, but scant attention has been given to it. To have considered carefully the role of the private colleges, however, would have further complicated an already knotty matter.

Efficient Pricing

Before embarking on our detailed investigation of equity, we turn to a brief analysis of some implications of seeking efficient pricing of higher education. As already noted, both kinds of considerations are relevant to the evaluation of alternative methods of financing public higher education.

The cost of a college education to a student and his family - apart from the income foregone - can be analysed in two parts. One is what can be termed the "price" of the education - the tuition charge, the books and supplies, and so forth. The second is the "ease of financing" the price - that is, the availability and terms of loan funds and scholarships.

The level of the price of college education, and the ease of financing it, are jointly relevant to individuals' decisions. An apparently high tuition rate may be quite manageable if grants or scholarships are widely available or if loans can be obtained at sufficiently low interest rates. Similarly, even a total failure of scholarship programmes and capital markets to provide financing assistance can turn out to be inconsequential if the total price of education (including foregone income) is sufficiently low. Thus, there would seem to be trade-off possibilities between the price of education and the means of financing it - combinations among which any particular individual would be indifferent.

But considerations of public policy dictate that we go beyond an analysis of any individual's preferences to take account of all the resources used up in the process of satisfying those preferences. Thus, we are led to consider the questions of what is a socially efficient price of education, as well as what is a socially efficient set of finance terms, including an interest rate.

Economic efficiency may be said to exist in a market when the price of the good or service is equated with the marginal opportunity cost (value of the best alternative use) of the resources used to produce it,

and both are equated with the benefits from an additional unit of the good. Thus, given the distribution of income, the preferences of all individuals in society, and the technological production possibilities, the efficient price for any given unit of production (e. g. man-year) of higher education is the price which is equal to the marginal net social cost of providing that education and the marginal benefit received by the student. By net cost we mean the marginal cost of pruduction minus any marginal "external" benefits, that is benefits that are not captured by the individuals whose education produced them[1]. To the extent that such external benefits occur, the efficient price to charge students would be below the marginal cost of producing the education services[2].

This view of pricing clearly implies that society (taxpayers in general) should subsidize higher education as a matter of efficiency. Since some external benefits may be realized within local areas while others may be distributed more broadly, all levels of government - federal, state, and local - would presumably share in the costs. Insofar as the bulk of externalities accrue at the national level, in part because of population migration, this would argue for a reallocation of public financing of higher education away from state and local governments and to the federal government.

Public subsidies can take a variety of forms. In addition to "low" tuition rates, there are low-cost loans, income tax credits or deductions to parents, and outright cash grants to students, all of which can be equivalent to a tuition reduction. Any of these forms, and no doubt others as well, could be used to produce the desired public subsidy and, in turn, result in an efficient "price". The choice among them rests largely on an equity consideration - that is, the extent to which persons not in "need" would benefit[3].

There are some individuals who may be "qualified"[4] for college but who will not attend college because the combinations of available price and interest charges are "excessive" relative to their financial situation and to the strength of their desire to attend. The willingness to incur these costs is conditioned by factors including family income and wealth, family size, and parental health. Yet there appears to be a social consensus that these factors ought not to bar college attendance, so that "needy" individuals with the ability and motivation to benefit from college should go.

If compulsion is to be avoided, these barriers to college attendance could be offset in three general ways: a) incomes of such students and their families might be supplemented; and/or b) the price of college education for them could be reduced; and/or c) the interest rate applicable to their borrowing for college could be reduced.

One might argue that the judgment that a student "should" go to college, even though family circumstances would lead him not to go, represents an implicit social decision that his family's income is "too low". Thus, an increase in family income would seem to be called for. If the objective, however, is to make it possible for this student to attend college at a minimum cost to others, then the approach of giving to needy students cash transfers that are not restricted as to use, is likely to be inefficient ; very substantial transfers might be required before any of the additional money would be used for the student's higher education. A possible variant is to restrict the use of cash grants to higher education. But this alternative may be difficult to implement, since as a practical matter there are no means for preventing some of the grant money from going to families - even some of them with very low incomes - whose children would have gone to college anyway and who now, having received the grant, will be able

1. At the conceptual level, the possibility of external costs as well as benefits should be considered. It is not generally argued, however, that such costs are notable, if, indeed, they occur at all in higher education.
2. It is difficult to estimate marginal costs, but it might reasonably be assumed that long-run marginal cost can be approximated by average instructional plus capital costs.
3. For further discussion of this issue in the context of income redistributional programmes see B.A. Weisbrod (JEC paper, vertical efficiency).
4. It seems clear that there is considerable arbitrariness in deciding who is "qualified" to benefit from a college education.

to increase their expenditures on other goods and services. Grants to such families are not necessarily undesirable, but the point is that grants may not be required to achieve educational objectives, however justified they may be from the point of view of a more general anti-poverty effort.

Consider now the alternatives of reducing the price and/or interest rate for the "needy". If, to begin with, the price and interest rate were set at levels that were economically efficient - in terms of the costs involved, as discussed in the preceding section - then further reductions would sacrifice some allocative efficiency in order to bring about effects that were deemed more equitable. Such a trade-off of efficiency for equity is by no means unique to higher education, nor is it necessarily undesirable.

In practice, each of the alternatives is bound to fall short of fully realizing equity objectives. Subsidies, whether in the form of cash, tuition rate reductions, or reductions in interest rates, are certain to go to some persons other than those whom "society" specifically wishes to assist, since the "needy" and "deserving" are frequently difficult to identify. Thus, subsidies go, at least to some extent, to the "wrong" people - with taxpayers, some of whom are themselves worthy of help, paying the cost.

Some perspectives on the dimensions of need can be obtained by a theoretical disaggregation of the population into several different groups. Group I includes those students (and their families) who are willing and able to pay at least the full long-run marginal cost (which we suggested above might be approximated by average instructional plus capital cost) net of estimated external benefits, and the full market interest rate. A portion of this group, while willing to pay these costs, can do so only by incurring some "hardship". Group II includes those who are willing and able to pay some lower, positive price and interest rate, and some fraction of this group could pay these amounts only with some hardship. Finally, Group III includes those people who would need bribes to cause them to attend college, being unwilling to attend at any combination of a positive price and positive interest rate. All three groups are defined to include only those deemed "eligible" - in terms of aptitude and motivation - to attend college.

One of the implications of the structuring of these three groups is that the amount of subsidy required to cause an individual to attend college is a continuous variable with a wide range of values. Some students will require very substantial subsidies and others none at all in order to provide full equality of opportunity in higher education.

Identifying those who are deserving of additional subsidies to enable them to go to college or to go without undue hardship is a most difficult task. Assume, however, that the "need" for higher education subsidies can be estimated in a satisfactory, if rough, manner, perhaps applying the standards used in student financial aid analyses. The perplexing question then is who should pay for these equity-based subsidies? Utilizing taxpayers in general as a source of revenue, while having merit, does imply that any sum of money that students and their families "cannot afford" to pay, can be paid by, and should be paid by taxpayers. But when it is borne in mind that "taxpayers in general" include many quite low income taxpayers, it becomes clear that a shifting of the financial burden from students and their families to taxpayers involves, to some extent, a shift of the burden to families whose incomes and ability to pay may be less than the ability to pay on the part of students and their parents.

This raises a more fundamental issue of the meaning of "ability to pay". Just as standards have been established for determining how much a family can "afford" to pay for higher education, so might standards be established to determine how much a family could "afford" to pay in taxes. If such a study were done, it might well conclude that families of given size, given needs, and with incomes below some specified amount, could not afford to pay any taxes at all; nevertheless, we suspect that many such families are, in fact, actually paying taxes - and would be required to pay even more taxes if state support for higher education were increased.

Another possible source of subsidy funds for the needy is other college students and their parents. We noted above that there are some families, particularly in Group I, who are able and willing to pay more than the efficient price of education. If they were charged a higher price, the subsidies required for needy students could be obtained outside the tax system. This would amount to the use of classic

price discrimination, to charge what the traffic will bear. One might think of the resulting schedule of charges as reflecting a sliding-scale college payment plan, with the possibility of negative charges for the most needy.[1]

On the assumption that a choice can be made regarding the most appropriate subsidy device for achieving greater equity, there is still a larger issue concerning the propriety of limiting subsidies to those who choose college rather than some other means for enhancing individual and social well-being. For the many young people not deemed qualified for college or not interested in attending college under any reasonable pricing conditions, there is a variety of other methods by which they can enhance their incomes and future satisfaction, and otherwise become effective citizens. Job training and investments in small businesses are only two substitutes to college-going. Whether from the standpoint of achieving equity or efficiency in resource allocation, it would be highly desirable to make these and perhaps other alternatives available to those young people who do not opt for college. A broadened subsidy programme might well be more costly. But it would at the same time do much to provide greater equality of opportunity for all young people, not merely for college students.

The relevance of the proposal for broadening the subsidy base will become clearer as we turn to the empirical results of our investigation of how the benefits and costs of public higher education are shared in California. We see that a large percentage of all young people receive no public subsidy at all through the public higher education system, while a small percentage receives very substantial public subsidies.

1. Michigan State University has been experimenting with such a plan, though in a quite restricted form. The sliding-scale approach is also implicit in cases where the size of scholarships is a function of "need".

II

THE DISTRIBUTION OF COSTS AND DIRECT BENEFITS OF PUBLIC HIGHER EDUCATION: THE CASE OF CALIFORNIA

The public higher education system in the United States provides - or, at least, offers - a public subsidy to young people of college age. The extent to which the young people actually receive the subsidies depends on a) whether they can qualify for admission, b) whether they avail themselves of the opportunity to attend, and, if they do, c) what quantity and quality of education they receive. As a result, the amount of subsidy received through the public financing of higher education varies greatly from one person to another. Our objective in the remainder of this paper is to estimate a) the amounts of subsidies received through higher education, b) the variation in subsidies received by students depending upon the amount of schooling and the kind of schooling they obtain, and c) the extent to which these subsidies are received in different amounts by students whose families are at different socio-economic levels.[1]

Attention is restricted to undergraduate education, and the data used are for public education in California. While higher educational systems differ among states, it would appear that the results for California are broadly characteristic of those for a number of other states.

A knowledge of the magnitude and distribution of subsidies or direct benefits provided through public higher education, or indeed through any public programme, is important for what it suggests as to appropriate pricing, tax, and expenditure policy. By "appropriate" we mean policies that will be efficient in the sense of doing the most to raise output, and at the same time equitable in the sense of doing the most to achieve society's distribution goals, such as providing greater equality of opportunity for young people. We can illustrate some of the possibilities. For those "eligible" for higher education, uniform subsidies may provide a "windfall" to the more financially able, while doing little to facilitate college attendance by the less well-off. This might argue for some kind of flexible pricing system in higher education, though much the same effect might be achieved less directly through the tax system. For those not eligible for public higher education, the provision of other kinds of subsidies or direct-benefit programmes may not only yield substantial benefits to others but also help to achieve greater equality - of both educational opportunity and of opportunity in general.

Subsidies Students Can Receive

The amounts of public higher education subsidies that college students can and do receive are the difference between tuition and the costs - instructional and capital - of providing instruction to them. The size of this difference for any student depends on the number of years of instruction received, and the subsidy per year of schooling. The latter, in turn, depends essentially on the costs of the particular college, and on its price (primarily tuition).

1. Little effort seems to have been given to this subject. For one interesting and perceptive foray, see Christopher Jencks, "Social Stratification and Mass Higher Education", Harvard Education Review, Spring 1968.

In 1965 the public subsidy provided through higher education in California ranged from $720 for a year at a Junior College to $1,350 and $1,450 for a year in the lower division (first two years) at a California State College and at the University of California, respectively. But the one-year subsidies tell only a portion of the subsidy story, for, while some students may attend a public college for only a year or even less, others attend for four years or more. And not only do those who attend for longer periods receive larger subsidies for that reason alone, but also because the subsidies increase as students progress to the upper division levels. For California, students who complete a two-year Junior College programme receive an average subsidy totalling $1,440, while those completing a baccalaureate programme at a State College receive four times as large a subsidy - $5,800 - and graduates from a University of California campus receive a four-year subsidy of more than $7,100. The actual amounts of subsidies vary, depending upon patterns of transfer among these three segments of the California public higher education system.

The proportions of students entering each segment of higher education who actually complete it vary considerably, from about 60% at the University, to 55% at the State Colleges, and to 30% at the Junior Colleges. But even this is deceptive, since many eligible students do not avail themselves of any public higher education. Some prefer to enter the work force, others enter the military service, and many females marry and do not continue their schooling. Still others enrol in private institutions of higher education in California, while another but smaller group seeks higher education outside of California.

Of those who do enrol in public higher education in California, the proportions eligible for each segment who actually enrol in that segment are often very low. For example, of the 19% of high school students eligible for the University of California in 1965, only 5% planned to enrol at the University; another 4% planned on going to State Colleges, 5% to Junior Colleges, 3% to other institutions, and 2% planned no further education. Of those 17% eligible for State Colleges (but not for the University), 2% planned to enrol at a State College, 8% at the Junior Colleges, and 4% did not plan to enrol at all. And with respect to the Junior Colleges, for which all students are eligible, only some 30% of high school graduates planned to enrol; this constituted one half of the 64% of high school graduates who were not eligible (on the basis of scholastic performance in high school) for either the University or a State College. Thus, whatever their reasons, many high school students enrol at public institutions of higher education in California which provide subsidies smaller than those for which they are eligible.

Just as the amount of public subsidy varies among the three segments of the California higher education system, so do the attrition rates. Students who enter a Junior College not only receive the smallest subsidy per year, but they are most likely to remain in school for only a short time. By contrast, students who enter the University of California receive a far greater subsidy per year and are most likely to receive that subsidy for four years, until graduation. The high attrition rate at the Junior College level reflects in part the fact that a number of its programmes require only one year of schooling. The rate of attrition at the State Colleges is somewhat lower, and attrition at the University of California is the lowest, largely as a result of its greater selectivity in admissions.[1] Its first-year attrition rate - 15% - seems rather high, but the four-year completion rate of 55% is within the range for most other comparable four-year institutions. However, an additional 3% of the initial entrants to the University of California completed their work at a State College, and some others undoubtedly graduated from colleges outside the California system of public higher education.

Distribution of Amounts of Subsidies

We have constructed a rough distribution of the percentage of an age cohort of high school graduates who receive different amounts of public subsidies for higher education, utilizing data on instructional and capital costs, transfer patterns among the three systems, and attrition rates. This information is summarized in Table 1. The rather startling conclusion is that while a small proportion - 9% - receives rather large subsidies, exceeding $5,000, more than half of California's young people receive under

1. For additional details see Benefits, Costs, and Finance of Higher Education, op. cit., Chapter 4.

Table 1. ESTIMATED DISTRIBUTION OF PUBLIC SUBSIDIES FOR HIGHER EDUCATION, BASED ON AMOUNT RECEIVED DURING PERIOD ENROLLED

AMOUNT OF SUBSIDY ($)	PERCENTAGE OF PERSONS RECEIVING
0	41
1 - 749	14
750 - 1,999	30
2,000 - 3,499	3
3,500 - 4,999	3
5,000 - 6,499	6
6,500 +	3
	100

$750 in total subsidy for higher education. And a substantial fraction - 41% - receive no subsidy at all. This group is divided between those who obtain no higher education whatsoever - almost 80% - and those who plan to attend private colleges within California or colleges outside the state - about 20%.

In short, there is a highly unequal distribution in the amounts of public subsidies actually received, even though California prides itself on the wide access to higher education it provides and the high enrolment ratios which are presumably a reflection of this. Moreover, there is little reason to believe that the distribution of public subsidies through higher education is less unequal in other states than it is in California. No state has as widely accessible a Junior College system as does California; thus, other states probably have larger proportions of young people who obtain little or no college education.

Distribution of Subsidies by Family Income

What can be said about the distribution of subsidies provided through higher education when measured against students' family income levels? While this is a difficult question to answer with the available data, we have tried to shed light on it.

To begin with, it is useful to examine the patterns of college-going by level of family income. These are shown in Table 2, where columns (3)-(6) show the family income distributions for all California public college students in 1964; column (2) shows the income distribution for families without children in California public higher education, and column (1) shows the distribution for all California families.

The distributions by family income clearly differ among the groups shown. Median family incomes (see bottom row of table) are highest for parents of university students, followed by State College student families and Junior College student families. Lowest of all is the median for all families without children in the California system. (This is heavily weighted with elderly and, on average, low income families.) These patterns are about what one might expect and, in general, conform to the patterns shown in other surveys.[1] Thus we conclude that access to subsidies is positively related to levels of

1. For example, see the Wisconsin data in L.J. Lins, A.P. Abell, and D.R. Stucki, Costs of Attendance and Income of Madison Campus Students, The University of Wisconsin, 1964-1965 Academic Year, Office of Institutional Studies, January 1967; I.M. Boyak, A.P. Abell, and L.J. Lins, Costs of Attendance and Income of University of Wisconsin-Milwaukee Students, 1964-1965 Academic Year, Office of Institutional Studies, March 1967; and L.J. Lins, A.P. Abell, and R. Hammes, Costs of Attendance and Income of University of Wisconsin Center Students, 1964-1966 Academic Year, Office of Institutional Studies, May 1966.

Table 2. DISTRIBUTIONS OF FAMILIES BY INCOME LEVEL AND TYPE OF COLLEGE OR UNIVERSITY, CALIFORNIA, 1964

(In percentage)

INCOME CLASS $	ALL FAMILIES	FAMILIES WITHOUT CHILDREN IN CALIFORNIA PUBLIC HIGHER EDUCATION	FAMILIES WITH CHILDREN IN CALIFORNIA PUBLIC HIGHER EDUCATION			
			TOTAL	JC	SC	UC
	(1)	(2)	(3)	(4)	(5)	(6)
0 - 3,999 ...	16.1	17.0	6.6	8.1	4.1	5.0
4,000 - 5,999 ...	14.8	14.9	13.0	15.9	10.2	7.5
6,000 - 7,999 ...	18.9	19.0	17.6	19.6	17.0	11.1
8,000 - 9,999 ...	18.1	18.3	16.4	16.9	17.2	13.1
10,000 - 11,999 ...	12.4	12.1	15.8	14.4	19.9	13.3
12,000 - 13,999 ...	7.4	7.3	8.8	17.2	10.8	11.3
14,000 - 19,999 ...	7.9	7.5	13.0	11.1	13.0	20.3
20,000 - 24,999 ...	1.8	1.6	3.4	2.6	3.3	6.6
25,000 + ...	2.6	2.3	5.4	4.2	4.5	11.8
Total	100.0	100.0	100.0	100.0	100.0	100.0
Median Income	$8,000	$7,900	$9,560	$8,800	$10,000	$12,000

NOTE: JC (Junior Colleges); SC (State Colleges); UC (University of California).

SOURCE: Column (1) - Letter from Office of Legislative Analyst, California Legislature, in Tuition for California's Public Institutions of Higher Education, Joint Committee on Higher Education, Hearings, 13th October, 16th October, 1967; see Tab T, Table 1.
Column (2) - Percentage distribution of Column (2), calculated by authors.
Column (3) - Weighted average of Columns (4), (5) and (6).
Columns (4), (5), (6) - Edward Sanders and Hans Palmer, The Financial Barrier to Higher Education in California (Claremont: Pomona College, 1965), Table M, p. 21, which relates to distribution of parent-supported students only.

family income, with the highest single-year subsidy going to UC students (and their families) who already have the highest median family incomes ($12,000).[1]

We can present some crude figures to illustrate the association of family income and subsidies received, by comparing median family incomes for the groups shown in Table 2 with the amounts of the subsidies going to each of these groups. Table 3 presents our estimates of these data. Median income of families of various types is shown in line 1, the one-year subsidy received is given in line 2a, and the subsidy as a percentage of family income is presented in line 2b. Because students first enrolling at each type of institution do not remain in college equally long, the average number of years they are enrolled is also shown, in line 3. The total subsidy received is shown in line 4a, and the percentage of family income that the subsidy constitutes is in line 4b. Because students transfer among the three higher education systems, the average subsidy is not simply the product of the average subsidy in a particular system and the average number of years of schooling obtained by students who begin their schooling in that system. As indicated by line 2b, the values of the single year subsidies vary from

1. Were we to relate the data shown in Table 2 to the data on subsidies received over the entire college stay, the differences in the subsidies received would be accentuated. The reason is that University of California students are more likely to complete four years than are State College students, and the latter are more likely to complete four years than the vast bulk of the students who begin at Junior Colleges.

Table 3. AVERAGE FAMILY INCOMES AND AVERAGE HIGHER EDUCATION SUBSIDIES RECEIVED BY FAMILIES, BY TYPE OF INSTITUTION CHILDREN ATTEND, CALIFORNIA, 1964

	ALL FAMILIES	FAMILIES WITHOUT CHILDREN IN CALIFORNIA PUBLIC HIGHER EDUCATION	FAMILIES WITH CHILDREN IN CALIFORNIA PUBLIC HIGHER EDUCATION			
			TOTAL	JC	SC	UC
	(1)	(2)	(3)	(4)	(5)	(6)
Average family income (i)	8,000	7,900	9,560	8,800	10,000	12,000
Average higher education subsidy per year (ii)						
a) Amount in dollars	-	0	880	720	1,400	1,700
b) Per cent of line 1	-	0	9	8	14	13
Average number of years of higher education completed	n. a.	n. a.	n. a.	1. 2	2. 6	2. 8
Average total higher education subsidy						
a) Amount in dollars	-	0	1,700	1,050	3,810	4,870
b) Per cent of line 1	-	0	18	12	31	41

(i) Median incomes from Table 2.
(ii) Average subsidies are based on the distribution of enrolment by year of school and on distribution of enrolment by type of institution.

zero per cent of family income for those without children in public colleges and universities (some of these people may have children in private colleges or in public colleges not in California), to 14% of family income for those families with State College students.

The average overall subsidy is equal to 9% of current money income for all parents of publicly-enrolled college students (line 2b, column 3), but the subsidy climbs to 18% of family income when we take account of the number of years that the educational subsidy is received (line 4b, column 3). Because, as noted before, the amount of schooling received differs, the average total subsidies (line 4a) rise far more sharply than the single-year subsidies (line 2a), as we contrast the families with children enrolled in California Junior Colleges, State Colleges, and University. These patterns of subsidies raise serious questions about the equity of the current system for financing public higher education in California.

At the same time, however, the distributions of students by parental income (as shown by each of the columns in Table 2) are so wide for each type of system - University of California, State College, and Junior College - that any strong conclusions about the "class-serving" nature of the entire system of higher education in California cannot be drawn. While there is a tendency for the higher subsidy schools to draw a higher-income clientele, the overlap of the distributions is still very substantial.

Some added light can be thrown on the equity issue by a restructuring of recent data presented by the California Co-ordinating Council for Higher Education.[1] The data from several of its tables have been

1. Co-ordinating Council for Higher Education, State of California, Financial Assistance Programs, 67-13 (Second Revision) 31st October, 1967, Table I-2, p. I-9; Table I-3, p. I-10; and Appendix Table B-3.

combined to show how eligibility and plans for higher education enrolment vary systematically with income.

We see in Table 4 that under 20% of high school graduates qualify for the substantial University subsidies; this is a product of the academic entry requirements. Even more striking is the fact that the percentage of all students qualifying for the University of California (column 1) rises quite dramatically by family income level - from about 10% in the lowest income bracket (under $4,000) to 40% in the highest (over $25,000). Thus the correlation between high school achievement and family income - and all that it reflects - is startling indeed. This pattern persists as we widen our view to include those eligible for both the University and those who are eligible for State Colleges (column 2). But a close examination of the differences between the two columns shows that the percentage of those eligible only for the State College system is roughly constant with respect to income level; thus, University eligibility requirements account largely for the unequal distribution of opportunity.

Table 4. DISTRIBUTION OF HIGH SCHOOL GRADUATES BY ELIGIBILITY FOR PUBLIC HIGHER EDUCATION IN CALIFORNIA, BY TYPE OF EDUCATION AND FAMILY INCOME

(In percentage)

FAMILY INCOME $	PERCENTAGE DISTRIBUTION OF HIGH SCHOOL GRADUATES BY ELIGIBILITY FOR	
	UNIVERSITY OF CALIFORNIA	UNIVERSITY OF CALIFORNIA AND STATE COLLEGES
0 - 3,999	10.7	28.0
4,000 - 5,999	11.5	26.3
6,000 - 7,999	11.9	30.5
8,000 - 9,999	16.2	33.2
10,000 - 12,499	19.4	37.1
12,500 - 14,999	22.5	39.8
15,000 - 17,499	27.9	45.4
17,500 - 19,999	29.5	45.1
20,000 - 24,999	33.3	46.1
25,000 +	40.1	54.3
Not reported	13.3	28.0
All	19.6	36.3

NOTE: Excluded from the sample of 8,162 were 302 students planning vocational training, 38 non-respondents on enrolment plans, and 20 for whom eligibility was indeterminate.

SOURCE: Based on data from CCHE, Financial Assistance Programs, 67-13, (Second Revision) 31st October, 1967, Table I-2, p. I-9; Table I-3, p. I-10; and Appendix Table B-3.

The extent to which family income influences the likelihood that a student who is eligible for a high subsidy school will go to it is indicated in Table 5. For the University (column 1) a larger fraction of upper than lower income students plan to attend; the same holds for the combined University-State College system group (column 2); and the pattern continues - though in somewhat muted fashion - when we consider all high school graduates (column 3). Actually, these results are somewhat deceptive since those eligible for a "higher" system can also attend a "lower" system. Indeed, when we compare the percentage of University eligible students planning to attend one of the three public systems, we find that the proportion is fairly constant with respect to family income, at about 70-75% (these data are not shown in the accompanying tables). Much the same kind of pattern emerges for both the University and

Table 5. COLLEGE ATTENDANCE PLANS OF CALIFORNIA HIGH SCHOOL GRADUATES, BY FAMILY INCOME AND HIGHER EDUCATION SEGMENT, 1966

(In percentage)

FAMILY INCOME LEVEL ($)	PER CENT OF UC ELIGIBLES PLANNING TO ATTEND UC	PER CENT OF UC-SC ELIGIBLES PLANNING TO ATTEND EITHER UC OR SC	PER CENT OF ALL CALIFORNIA HIGH SCHOOL GRADUATES PLANNING TO ENROL IN UC, SC OR JC
0 - 3,999	30.4	22.5	53.1
4,000 - 5,999	26.1	29.7	56.1
6,000 - 7,999	23.4	28.1	56.3
8,000 - 9,999	21.5	36.5	60.0
10,000 - 12,499	25.3	32.6	62.0
12,500 - 14,999	26.2	37.5	64.6
15,000 - 17,499	26.9	32.1	63.4
17,500 - 19,999	33.3	45.7	64.2
20,000 - 24,999	45.4	52.0	68.2
25,000 + ...	46.7	47.8	57.8
No response	30.5	30.1	47.9

SOURCE: Same as Table 4.

State College eligibles who plan to undertake higher education. The point, however, is that enrolment in a lower system - often dictated by family income considerations - implies a reduced level of subsidies.

Who Pays the Taxes?

Having shown the extent to which families in different income groups are awarded subsidies through the fiscal system by virtue of the provision of higher education, we turn now to the question of how these subsidies are financed. Specifically, we estimate distributions of state and local taxes paid by families at each income level. The objective is to provide a basis for comparing the subsidies received with the tax payments made. Such information is essential in assessing the equity of the current methods of financing higher education in the State of California.

Our approach is to estimate the incidence of the most important state and local taxes by family income level, so as to note the absolute amount of taxes paid at each income level. We can then compare this amount with the subsidy received and note any differences. But we still have no real way of determining how much of whatever taxes are paid reflect support for higher education, as against the many other services provided by state and local governments.

The average amount of taxes paid at each income level as well as the effective tax rate, for California state taxes alone, and for state and local taxes combined, is shown in Table 6. The most important finding is that while the state tax structure (column 2) seems to be somewhat progressive - that is, the effective tax rate rises with income - except in the lowest income classes, the combined state and local tax structure (column 4) is regressive below $8,000 and is essentially proportional above that level.[1]

1. The recent, 1967, changes in the California state income tax structure have increased, but only slightly, the overall progressivity of the state tax structure.

Table 6. ESTIMATED TAX BURDENS BY INCOME CLASS, CALIFORNIA, 1965

ADJUSTED GROSS INCOME CLASS ($)	STATE TAXES ONLY PER FAMILY a) ($)	EFFECTIVE STATE TAX RATE b) (%)	STATE AND LOCAL TAXES PER FAMILY c) ($)	EFFECTIVE STATE AND LOCAL TAX RATE b) (%)
	(1)	(2)	(3)	(4)
0 - 3,999	104	5.2	474	23.7
4,000 - 5,999	132	2.6	527	10.5
6,000 - 7,999	161	2.3	576	8.2
8,000 - 9,999	221	2.4	696	7.7
10,000 - 11,999	301	2.7	833	7.6
12,000 - 13,999	389	3.0	984	7.6
14,000 - 19,999	539	3.2	1,228	7.2
20,000 - 24,999	865	3.8	1,758	7.8
25,000 + ...	2,767	5.5	4,093	8.2

SOURCES: Personal income, sales, cigarette and beverage taxes by income level were obtained from Letter from Office of Legislative Analyst, State of California in Tuition for California's Public Institutions of Higher Education, Joint Committee on Higher Education, Hearings, 13th and 16th October, 1967; see Tab. T, Table 1. State gasoline taxes and local property taxes were based on itemised tax deductions reported on state income tax returns, 1965, and summarized in Franchise Tax Board, Annual Reports, 1965 and 1966, Table 13. Local sales taxes were assumed to be distributed in the same manner as state sales taxes above. Since local sales tax revenues in 1965 equalled one third of state sales tax revenues, this factor was applied to the estimated amount of state sales taxes in each income level.

a) Personal income, state sales, cigarette, and alcoholic beverage taxes only.

b) Taxes as a per cent of estimated mean income of each income class. The mean of the highest income interval was arbitrarily assumed to be $50,000.

c) State taxes include personal income, sales, cigarette, alcoholic beverage, and gasoline taxes. Local taxes include local sales and property taxes.

We return now to the major task of this section - to compare the taxes paid with the subsidies received by families with children enrolled in college, so that we can observe the extent to which broad groups of families do or do not receive net subsidies through higher education. In making such comparisons we once again remind the reader that this involves comparing all taxes with benefits received from higher education alone. As shown by Table 7, the annual value of higher education subsidies (line 2) received by a family with a single child enrolled in a public college exceeds the total amount of all state and local taxes they pay (line 3), by rather substantial amounts. On an overall basis the average higher education subsidy is $880 per year (line 2, column 3), in contrast to total state and local taxes paid of $740 (line 3, column 3); this results in an annual net transfer of $140 from all taxpayers to parents of each college student. But this average conceals wide differences by type of college.

For families with a child at one of the State Colleges or one of the University campuses, the net transfers range from $630 to $790 per year. Meanwhile, families without children or with children not enrolled in public institutions of higher education receive no subsidy whatsoever, while they pay an average of $650 in state and local taxes. This is not to suggest that such families should pay no state and local taxes, for some may have benefited in the past, others may benefit in the future, and still others may have opted for more expensive non-public California higher education. Moreover, state and local taxes finance public services other than higher education. In any case, as is evident from a comparison of line 4 and line 1, the current method of financing public higher education leads to a redistribution of income from lower to higher income families; indeed, there is very substantial progressivity in the resulting pattern of transfers.

Table 7. AVERAGE FAMILY INCOMES, AVERAGE HIGHER EDUCATION SUBSIDIES RECEIVED, AND AVERAGE STATE AND LOCAL TAXES PAID BY FAMILIES, BY TYPE OF INSTITUTION CHILDREN ATTEND IN CALIFORNIA, 1964

	ALL FAMILIES	FAMILIES WITHOUT CHILDREN IN CALIFORNIA PUBLIC HIGHER EDUCATION	FAMILIES WITH CHILDREN IN CALIFORNIA PUBLIC HIGHER EDUCATION			
			TOTAL	JC	SC	UC
	(1)	(2)	(3)	(4)	(5)	(6)
1. Average family income a)	8,000	7,900	9,560	8,800	10,000	12,000
2. Average higher education subsidy per year b)	-	0	880	720	1,400	1,700
3. Average total state and local taxes paid c)	620	650	740	680	770	910
4. Net transfer (line 2 - line 3)	-	-650	+140	+40	+630	+790

a) From Table 2.
b) From Table 3.
c) Total state and local tax rates from Table 9 were applied to the median incomes for families in each column.

CONCLUSION

Public policy regarding higher education must consider a number of factors among which the economic efficiency of expenditures on higher education and the distributional equity of the public support for higher education are surely prominent. After a brief analysis of the economic efficiency issue, this paper turned to its primary objective - an empirical investigation of the distributional effects of public higher education in our most populous state, California.

The general nature of the redistributive effects of the current method of financing public higher education in California is clear. Some low income persons have benefited handsomely from the availability of publicly-subsidized higher education. But on the whole, the effect of these subsidies is to promote greater rather than less inequality among people of various social and economic backgrounds, by making available substantial subsidies that lower income families are either not eligible for or cannot make use of because of other conditions and constraints associated with their income position.

To overcome the effects of the present system would require a substantial overhaul of the pricing system in public higher education, a realignment of the tax structure, and/or a broadening of the eligibility base for public expenditure programmes. With respect to the latter alternative, eligibility for public subsidies to young people might well be expanded so as to embrace all young people - not only those who go on to college but also those who opt for alternative ways of expanding their earning power, such as apprenticeship or on-the-job training, or even investments in businesses. In any case, it is clear that whatever the degree to which our current higher education programmes are rooted in the search for equality of opportunity, the results still leave much to be desired.