DOCUMENT RESUME

ED 105 106 CE 003 356

AUTHOR Kalcic, Dismas B.
TITLE Interarea Educational Earnings Differentials: A Cross-Section Analysis.
SPONS AGENCY Manpower Administration (DOL), Washington, D.C. Office of Research and Development.
REPORT NO DLMA-91-29-72-30-1
PUB DATE 3 Dec 74
NOTE 109p.; Ph.D. Dissertation, Washington University, St. Louis, Missouri

EDRS PRICE MF-$0.76 HC-$5.70 PLUS POSTAGE
DESCRIPTORS Cross Sectional Studies; Doctoral Theses; *Educational Background; Females; Labor Unions; Local Unions; Males; *Metropolitan Areas; Negroes; Racial Discrimination; Salaries; *Salary Differentials; *Sex Discrimination; Subemployment; Tables (Data); Unskilled Labor; *Wages

ABSTRACT

This study was designed to explain the earnings differentials between metropolitan areas for six labor types, identified by level of education: 0-7, 8, 9-11, 12, 13-15, and 16 or more. Three separate models relate the average earnings of these types to several area variables. Multiple regression equations, based on 1960 Census and related data, indicate that: (1) the earnings of the male labor types are significantly related to the cooperating labor force in an area; (2) area unionization provides a significant explanation for the interarea earnings differentials of all labor types, both male and female; (3) the earnings of the lowest labor type are related primarily to the extent of area unionization and to the percentage of the type relative to the area labor force; (4) the earnings of females are affected by the number of females relative to males in the labor force with nonwhite females of the lowest labor type suffering discrimination; and (5) a North-South regional differential exists for females but not for males. These results indicate that the forces which determine the earnings of women are different from those which determine the earnings of men. (Author)

ED105106

TITLE: INTERAREA EDUCATIONAL EARNINGS DIFFERENTIALS: A CROSS-SECTION ANALYSIS

GRANTEE: Washington University
St. Louis, Missouri 63130

AUTHOR: Dismas B. Kalcic
5601 College Road
Lisle, Illinois 60532

This report was prepared for the Manpower Administration, U.S. Department of Labor, under research and development grant No. 91-29-72-30. Since grantees conducting research and development projects under Government sponsorship are encouraged to express their own judgment freely, this report does not necessarily represent the official opinion or policy of the Department of Labor. The grantee is solely responsible for the contents of this report.

Abstract

INTERAREA EDUCATIONAL EARNINGS DIFFERENTIALS:
A CROSS-SECTION ANALYSIS

by Dismas Kalcic

Chairman: Charles Leven

This study is designed to explain the earnings differentials between metropolitan areas for six labor types, identified by years of education: 0-7, 8, 9-11, 12, 13-15, and 16 or more. Three separate models relate the average earnings of these types to several area variables. All three are developed on the assumption that every area can be described by the same aggregate linear homogeneous production function. The earnings of a labor type are related through that function to other productive factors in the area. Model I limits the analysis to the level of capital relative to the labor force and to the quantity and quality of cooperating labor. Model II allows for short-run disequilibrium by adding a demand variable to the first model. Model III introduces a sex and race variable to account for quality differences in the area labor force, and it adds the degree of unionization as an institutional force on labor type earnings. Subsequently, all three models are modified by the addition of a regional dummy variable (North = 0; South = 1) to account for possible regional variance.

Multiple regression equations, based on 1960 Census and related data, were developed to test the several models. The main conclusions can be summarized: 1) that the earnings of the male labor types in general are significantly related to the cooperating labor force in an area; 2) that area unionization

provides a significant explanation for the interarea earnings differentials of all labor types, male and female; 3) that the earnings of the lowest labor type are related primarily to the extent of area unionization and to the percentage of the type relative to the area labor force, and that the earnings of the females are affected by the number of females relative to males in the labor force with nonwhite females suffering the additional burden of discrimination; and 4) that a North-South differential exists for females but not for males. Of most importance, the study indicates that the forces which determine the earnings of women are clearly different from those which determine the earnings of men.

BIBLIOGRAPHIC DATA SHEET	1. Report No. DLMA 91-29-72-30-1	2.	3. Recipient's Accession No.
4. Title and Subtitle INTERAREA EDUCATIONAL EARNINGS DIFFERENTIALS: A CROSS-SECTION ANALYSIS			5. Report Date 3 December 1974 6.
7. Author(s) Dismas B. Kalcic			8. Performing Organization Rept. No.
9. Performing Organization Name and Address Department of Economics Washington University St. Louis, Missouri 63130			10. Project/Task/Work Unit No. 11. Contract/Grant No. DL 91-29-72-30
12. Sponsoring Organization Name and Address U.S. Department of Labor Manpower Administration Office of Manpower Research and Development 601 D Street, N.W., Washington, D.C. 20213			13. Type of Report & Period Covered Final 4/1/72 to 8/30/74 14.

15. Supplementary Notes

16. Abstracts This study was designed to explain the earnings differentials between metropolitan areas for six labor types, identified by level of education: 0-7, 8, 9-11, 12, 13-15, and 16 or more. Multiple regression equations, based on 1960 Census and related data, indicate that: 1) the earnings of the male labor types are significantly related to the cooperating labor force in an area; 2) that area unionization provides a significant explanation for the interarea earnings differentials of all labor types, both male and female; 3) that the earnings of the lowest labor type are related primarily to the extent of area unionization and to the percentage of the type relative to the area labor force; 4) that the earnings of females are affected by the number of females relative to males in the labor force with nonwhite females of the lowest labor type suffering discrimination; and 5) that a North-South regional differentials exists for females but not for males. These results indicate that the forces which determine the earnings of women are different from those which determine the earnings of men.

17. Key Words and Document Analysis. 17a. Descriptors

Earnings, Education, Females, Labor, Labor unions, Males, Manpower, Negroes, Professional Personnel, Regional Planning, Skilled workers, Unionization, Unskilled workers, Urban areas, Urban planning

COSATI Field Group: I

17b. Identifiers/Open-Ended Terms

17c. COSATI Field/Group I

18. Availability Statement Distribution is unlimited. Available from National Technical Information Service, Springfield, Va. 22151.	19. Security Class (This Report) UNCLASSIFIED	21. No. of Pages 107
	20. Security Class (This Page) UNCLASSIFIED	22. Price

WASHINGTON UNIVERSITY

Department of Economics

Dissertation Committee:

Charles Leven, Chairman
Edward Kalachek
Fredric Raines

INTERAREA EDUCATIONAL EARNINGS DIFFERENTIALS: A CROSS-SECTION ANALYSIS

by

Dismas Bertrand Kalcic

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

November, 1974

Saint Louis, Missouri

ACKNOWLEDGMENTS

This dissertation marks the conclusion of an unusually long academic struggle. The patience and understanding of both the committee members and my Procopian brothers was not deserved, and for that reason all the more appreciated. Professor Charles Leven provided the original idea in suggesting that earnings and area resources are related in a consistent way. He also assisted in various capacities from beginning to end. Professor Edward Kalachek gave freely of his time to direct, assist, and encourage. Professor Fredric Raines was extremely gracious in consenting to serve on the committee after the dissertation was already well developed.

Bernard Glos and Warren Farb assisted with the computer program and criticized early drafts. Mrs. Fran Leitherer, Mrs. Anne Veverka, and my sister, Diane, provided enthusiastic clerical assistance.

The material in this project was prepared under Grant No. 91-29-72-30 from the Manpower Administration, U.S. Department of Labor, under the authority of Title I of the Manpower Development and Training Act of 1962, as amended. Researchers undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment. Therefore, points of view or opinions stated in this document do not necessarily represent the official position or policy of the Department of Labor.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

The existence of wage or earnings differentials between different geographic locations for work performed with comparable skill has been a subject of continuing interest among economists for both practical and scientific reasons. The scientific interest can be viewed as a challenge intimated in a comment made by Paul A. Samuelson over two decades ago:

> ... I fear that when the economic theorist turns to the general problem of wage determination and labor economics, his voice becomes muted and his speech halting. If he is honest with himself, he must confess to a tremendous amount of uncertainty and self-doubt concerning even the most basic and elementary parts of the subject.[1]

Many efforts have since been made to reduce this judgment to the level of an interesting anachronism, but even though they added insight to the problem, the best efforts have suggested that the forces of wage determination will not be easy to describe and that an empirically validated theory or system of theories will be long in coming.[2]

---

[1]Paul A. Samuelson, "Economic Theory and Wages," in David McCord Wright, ed., The Impact of the Union, New York, Harcourt, Brace, and Company, 1951, p. 312. In reference to Samuelson's comment, Campbell R. McConnell suggests that "a careful evaluation of the field today would yield a similar conclusion." (Campbell R. McConnell, Perspectives on Wage Determination, New York, McGraw-Hill Book Company, 1970, p.1).

[2]J.W.F. Rowe seems to have been something of a prophet when he wrote: "... it seems questionable in fact whether there can be one all-embracing theory of wages, which will sufficiently satisfy our sense of reality, and stand the test of historical experience." (J.W.F. Rowe, Wages in Practice and Theory, Routledge Press, London, 1928, p. 192.) Rowe seems to have grasped at a rather early time in the development of labor economics the continuing complexity of wage analysis.

The practical interest in geographic differentials can be attributed to more than one source. A verified differential can be the basis for both testing a theory and developing new theories,[3] and it can become both the basis for governmental policy decisions and the source of judging the success of those decisions in an examination of change over time.

The analyses to date have centered on differentials between broad economic regions, between states of the union and groups of states, and between metropolitan areas. Of these, metropolitan areas seem to be the most consistent with the idea of a local labor market, for any model defined over a broad geographic region such as a state or group of states, etc., almost certainly contains several separate labor markets. However, even with this apparent advantage, relatively few studies have concentrated on metropolitan areas.

In an early study, Edwin Mansfield focused on the relationship of income and city size and concluded that the average income of a defined consumer unit increases with city size.[4]

---

[3]This includes deriving quantitative estimates of important relationships, estimating production functions for individual industries, calculating elasticities of substitution between labor and capital, analyzing income distribution, population migration, changes in the location of manufacturing, etc.

[4]Edwin Mansfield, "City Size and Income, 1949," in Regional Income, Princeton University for National Bureau of Economic Research, 1957. The consumer unit is defined as "... a family or unrelated individual. A family is a group of two or more persons who reside together and are related by blood, marriage, or adoption. An unrelated individual ... is not living with relatives." p. 271.

More recently, Victor R. Fuchs sought to measure the influence of city size on the average hourly earnings of standardized groups such as white males, nonwhite males, etc. He found a strong and consistent positive relation between the earnings of these groups and the size of a city.[5] City size also explained one-third of the observed differential between southern and non-southern cities, with another third being explained by the labor force composition as measured by color, age, sex, and education.[6] Most recently, Philip R. P. Coelho and Moheb A. Ghali studied the economic influences on average money wages in several industries of a select group of metropolitan areas and found that the differentials could be fully explained by the relative cost of living.[7] Mark L. Ladenson challenged this finding, and even

---

[5]Victor R. Fuchs, Differentials in Hourly Earnings by Regions and City Size, 1959, National Bureau of Economic Research Occasional Paper 101, Columbia University Press, 1967.

[6]Studies previous to those by Mansfield and Fuchs examined the relationship between income and size of place, and the findings were similar. (See, e.g., Herbert E. Klarman, "A Statistical Study of Income Differences Among Communities," Studies in Income and Wealth, 6, National Bureau of Economic Research, New York, 1943; Milton Friedman and Simon Kuznets, Income from Independent Professional Practice, National Bureau of Economic Research, New York, 1945; D. Gale Johnson, "Some Effects of Region, Community, Size, Color, and Occupation on Family and Individual Income," Studies in Income and Wealth, 15, National Bureau of Economic Research, New York, 1952.) Herman P. Miller suggests that differences in occupational structure are the basis of the correlation between income and city size "since the largest centers contain a somewhat larger proportion of professional workers and a ... smaller proportion of laborers." See Herman P. Miller, Income of the American People, Wiley Publishers, New York, 1955, pp. 38-39.

[7]Philip R. P. Coelho and Moheb A. Ghali, "The End of the North-South Wage Differential," American Economic Review, 61, December 1971, pp. 932-937.

though Coelho and Ghali answered the challenge to some extent, the small sample for their analysis encourages great caution in the acceptance of their results.[8]

As indicated, these studies concentrated on the average income of a consumer unit, on the average hourly earnings of several standardized groups, and on the average wages of an industry. Although some insight was gained from these efforts, the implications that can be drawn for an individual laborer or a small group of laborers identified by occupation, skill group, or level of education, are highly limited. In one or more of the studies, earnings are related to qualitative characteristics of the labor force such as sex, race, education, etc., but none establishes a close theoretical relationship between the quality and quantity of cooperating factors in the productive process and the earnings of an individual or relatively homogeneous group.

The efforts in this study will be directed toward developing three models which will relate the average earnings of a labor type, as identified by level of education, to the level of capital relative to the labor force, to the quantity and quality of cooperating labor, and to the degree of unionization in the metropolitan area. The models will be tested with data listed in the 1960 Census of Population and related data.

---

[8]Mark L. Ladenson, "The End of the North-South Wage Differential: Comment," American Economic Review, 63, September 1973, pp. 754-756 and Philip R. P. Coelho and Moheb A. Ghali, "The End of the North-South Wage Differential: Reply," American Economic Review, 63, September 1973, pp. 757-762. The analysis was based on five northern and five southern cities.

1. Plan of the Study

A rationale for the models will be developed in Chapter Two. All three models will be built on the assumption that every area can be described by the same linear homogeneous production function. The differences in earnings of any labor type are related through that function to other productive factors in the area. Degree of area unionization is later introduced as an institutional force which may also account for differences in earnings.

Chapter Three describes the empirical test of the models, including a detailed list of data sources and formulas of computation for variables that are not published and which have to be estimated from published data. As suggested by the theory, multiple regression equations are developed for each of six educational labor types as distinguished by sex.

The fourth Chapter presents an econometric analysis of the area characteristics which determine the earnings of a labor type as indicated by the regression results. Although the theory is not completely supported in any one regression equation, it is generally supported in the male equations. A comparison of the several models warrants conclusions which are consistent with the results of previous studies and which offer new avenues for reflection. The final section of Chapter Four lists these conclusions and indicates some policy implications.

CHAPTER TWO

SOURCES OF EARNINGS DIFFERENTIALS BETWEEN METROPOLITAN AREAS

According to neoclassical theory, wages for a given type of labor should be identical in separate metropolitan areas under the conditions of perfect competition and long-run equilibrium. Where wage rates have not been equalized, conventional economic theory and previous research in the area of wage differentials and the economics of human capital suggest several major sources which could account for the lack of uniformity. There could be barriers to the free flow of resources among areas, which implies different ratios of capital to labor or of cooperating labor to a particular type of labor, and areas could experience different degrees of unionization or discrimination against females and nonwhites.

The ratio of cooperating labor to labor of a particular type is of special interest because it has been discussed for a long time in the theory of the firm, but attempts to measure the effect of cooperating factors on earnings have been avoided because of the "impossibility of measurement."[1]

---

[1]Interest in cooperating factors flows primarily from the marginal productivity theory. Williams' comment expresses what might be called a characteristic attitude of present-day labor economists toward the theory: "... the system adjusts constantly in a multitude of ways to changes in relative scarcities of factors of production. It is a recognition of this fact and of the fact that no other theory consistent with the results can

An area labor force can be distinguished according to level of human capital, sex, and race. Furthermore, this distinction can be made under the assumption that females and nonwhites are essentially different resources from males and whites, and not simply because females and nonwhites are generally paid less than their male and white counterparts. Women in general cannot perform at jobs requiring physical strength with the same skill and ease as men, and it has been suggested that nonwhites with a corresponding level of education have generally not received an education comparable to whites.[2] However, it may also be true that women and nonwhites have not received equal pay with males and whites for comparable jobs exercised with comparable skill. Thus, women and nonwhites may be paid less than males and whites either for the economic reason of lower productivity or for the institutional reason of discrimination by employers.

The importance of unions in effecting a wage differential is well established.[3] Their importance in effecting

---

explain so much that leads us to believe that the marginal productivity theory has real meaning." C. Glyn Williams, Labor Economics, John Wiley & Sons, Inc., New York, 1970, p. 130. The major limitation, according to Gitelman, is "its construction of propositions in such a form that they cannot be empirically validated." H. M. Gitelman, "An Investment Theory of Wages." Industrial and Labor Relations Review, 21, April 1968, p. 323.

[2]Much interest has been shown in this proposition. See, for example, Finis Welch, "Black-White Differences in Returns to Schooling," American Economic Review, 68, December 1973, pp. 893-907. This article contains a short bibliography of recent research.

[3]The most complete study to data is H. Gregg Lewis, Unionism and Relative Wages in the United States, University of ChicagoPress, 1963.

wage differentials between metropolitan areas can be reasonably concluded from the combined effect of "equitable comparison" and what has been referred to as "unionization spillover."

When a union bases its wage demands upon what other, comparable groups of workers are being paid, it is said to have made its decision on the basis of an "equitable comparison." Arthur M. Ross has considered this to be a major force in union negotiations:

> ... the most powerful influence linking together separate wage bargains into an interdependent system is the force of equitable comparison. This force does not spread evenly over the entire system, but runs in limited circuits. Even if big unions were broken up into little ones, as some advocate, equitable comparison would still have a pervasive effect.[4]

Equitable comparison can be considered as a link which ties together wages in the unionized sector of an area labor force.

There is evidence that contracts made in the unionized sector influence the working agreements of the nonunion labor force, that union influence "spills over" into the nonunionized sector.

In a cross-section analysis of interindustry wage differentials, Otto Eckstein and Thomas A. Wilson were unable to isolate the effects of unionism, and they conclude that they were unable to because

> ... any wage increase caused by unionization would permeate much of the rest of the wage structure in other industries and, therefore, would not appear in cross-section comparisons.[5]

---

[4]Arthur M. Ross, Trade Union Wage Policy, University of California Press, Berkeley, 1956, p. 8.

Robert Ozanne concurred with the reasoning of Eckstein and Wilson in his study of wage movements in the Chicago Plant of the McCormick Harvesting Machine Company (International Harvester Company since 1902) and chose to use time periods rather than interindustry comparisons because

> ... the wage causal variables appear to be more associated with time periods than with industries. ... [The] major phenomenon of average rate of wage change varying significantly over time is missed by the current emphasis on interindustry studies of both the cross-sectional (in the same year) and the over-time types. The major causal factors behind wage movements are apparently rarely confined to one industry but spill over and diffuse among many industries. There are good theoretical grounds for supposing that the wage effects of labor market conditions and trade unionism spread rapidly from firm to firm and industry to industry.[6]

The "good theoretical grounds" referred to by Ozanne very likely include equitable comparison on the part of unions and the tendency of nonunionized industries to pay as much as unionized industries both to be able to acquire high-quality labor and to keep their workers from organizing into a union. This tendency was substantiated in his comparison of wage movements and union activity in that one Chicago plant from 1860 to 1960. As analyzed by Ozanne, the McCormick and Chicago experience included at times big wage increases resulting from efforts to avoid unionism during the union organizing drive.[7]

---

[5]Otto Eckstein and Thomas A. Wilson, "The Determination of Money Wages in American Industry, "Quarterly Journal of Economics, 76, August 1962, p. 401.

[6]Robert Ozanne, Wages in Practice and Theory, University of Wisconsin Press, Madison, 1968, p. 11.

In a summary statement, Ozanne comments:

> It is highly unlikely that the experience of the McCormick works was unique. The union with which the McCormick Company bargained from 1862 through 1885 was a city-wide union, implying that the wage patterns observed at McCormick were city-wide patterns. The McCormick and Chicago experience was probably typical of industry located in urban centers nationally. On each occasion that unionism reappeared at the McCormick works -- 1903, 1916, 1919, the mid 1930's -- it was in the midst of great national surges in union growth.
>
> The large amount of unionism today with its heavy spillover effects makes it difficult to isolate the wage effect of unionism. The McCormick study, by permitting observation of varying amounts of unionism over time, aids in isolating this factor.[8]

If the earnings of a labor type in a metropolitan area are affected by a spillover, it is reasonable to assume that the strength of the spillover will vary with the degree of unionization. Under this assumption and the assumption that the earnings of all labor types within a metropolitan area are affected by the forces of equitable comparison and unionization spillover, the degree of unionization can be considered a major source of interarea differentials.

To place all the sources discussed above into a consistent and acceptable theoretical framework, three separate models will be developed. In the construction of these models, two basic assumptions will be shared by all three: 1) that all economic areas are described by the same aggregate linear homogeneous production function, and 2) that a labor type can be adequately identified by a given level of education.

---

[7] Ibid., pp. 64-65.

[8] Ibid., pp. 127-128.

For clarification of the latter, Griliches' comments are most germane:

> When we abandon the notion of one aggregate labor input and are faced with a list of eight major occupations, eight schooling classes, several regions, two sexes, at least two races, and an even longer list of detailed occupations, there doesn't seem to be much point in trying to distinguish all these aspects of the labor force simultaneously. The next small step is obviously not in the direction of a very large number of types of labor but rather toward the question of whether there are a few underlying relevant "dimensions" of "labor" which could explain, satisfactorily, the observed diversity in the wages paid to different "kinds" of labor. The obvious analogy here is to the hedonic or characteristics approach to the analysis of quality change in consumer goods, where an attempt is made to reduce the observed diversity of "models" to a smaller set of relevant characteristics such as size, power, durability, and so forth. One can identify the "human capital" approach as a one-dimensional version of such an approach. Each person is thought of as consisting of one unit of raw labor and some particular level of embodied human capital. Hence, the wage received by such a person can be viewed as the combination of the market price of "bodies" and the rental value of units of human capital attached to (embodied in) that body:
>
> $$W_i = w_o + rH_i + u_i$$
>
> where u_i stands for all other relevant characteristics (either included explicitly as variables, controlled by selecting an appropriate sub-class, or assumed to be random and hence uncorrelated with H_i).[9]

The human capital in Griliches' equation will be identified by six labor types, distinguished according to years of education: 0-7, 8, 9-11, 12, 13-15, and 16 or more. This distinction is encouraged by the 1960 Census data (to be used later) which reveal that the mean earnings for each of the

[9]Zvi Griliches, "Notes on the Role of Education in Production Functions and Growth Accounting," in Education, Income, and Human Capital, (edited by W. Lee Hansen), Columbia University Press, New York, 1970, pp. 89-90.

above categories was significantly different for 1959, the year for which the data were gathered: $3,658; $4,725; $5,500; $6,398; $7,846; and $10,863.[10]

The difference in these earnings is substantial and the causes are not apparent. It is not evident why one or two years of schooling should make so much difference in earning power -- why, for example, someone with a high school diploma is expected to earn approximately $900 more than someone who has completed only three years of high school. Herman P. Miller suggests two reasons: 1) that the income differentials between persons who attain a given level of schooling and those who graduate from that level reflects in part, differences in "ability," and 2) that lack of education limits opportunities for advancement to higher-income occupations.[11] The first gives strong support to using the six categories in order to identify different labor types. The second suggests that there is a correspondence between level of education and the requirements for a particular occupation.[12]

---

[10] Herman P. Miller, Income Distribution in the United States, U.S. Government Printing Office, Washington, 1966, p. 140.

[11] Ibid., pp. 143-144.

[12] Otis Duncan estimated the percentages of the total differences in incomes of persons with different levels of education which could be attributed to occupational differences and found that "well-educated persons engage disproportionately in high-income occupations and poorly-educated persons in low-income occupations." Otis D. Duncan, "Occupational Components of Educational Differences in Income," Journal of the American Statistical Association, 56, December 1961, p. 78. Anne

1. Model I

It was stated previously that geographic wage differentials may be partly explained by different factor proportions. A development of the first model, which assumes that this is the only reason for earnings differentials between metropolitan areas, can begin with the expression:

$$e_{ij} = f(KL_j, L_{1j}, \ldots , L_{24j}) \qquad \text{Eq. 2.1}$$

where

e = average annual earnings

KL = capital-labor ratio

L = percentage of the labor force having completed a specific level of education

subscript i = educational category

subscript j = metropolitan area

subscripts 1 to 6 = educational categories of white males

subscripts 7 to 12 = educational categories of white females

subscripts 13 to 18 = educational categories of nonwhite males

subscripts 19 to 24 = educational categories of nonwhite females

---

Mayhew also studied the relationship between occupation, earnings, and education. She found that her conclusions did not contradict Duncan for the population as a whole, but that education has different occupational and income effects for those who attend college and those who do not. Her analysis of the higher earnings of high school graduates led her to conclude that "they earn more in occupations which they would likely have been in even had they not stayed in school for the few additional years." Anne Mayhew, "Education, Occupation, and Earnings," Industrial and Labor Relations Review, 24, January 1971, pp. 224-225. This finding supports the formula to be used in the empirical portion of this study to calculate the average earnings of the various labor types as identified by level of education.

This equation is based upon the assumption that females are essentially different resources from males and nonwhites are essentially different resources from whites.

If the nonwhite males are in the same proportion to the white males for every educational category in every area, and if the nonwhite females are in the same proportion to the white females for every educational category in every area, the distinction of labor by race is irrelevant and Eq. 2.1 can be reduced to

$$e_{ij} = f(KL_j, L_{1j}, \dots , L_{12j}) \qquad \text{Eq. 2.2}$$

where $L_1, \dots , L_6$ are the educational categories of all males and $L_7, \dots , L_{12}$ are the educational categories of all females.[13]

Furthermore, if the females are in the same proportion to males for every educational category in every area, the distinction by sex is unnecessary and Eq. 2.1 can be further reduced to

$$e_{ij} = f(KL_j, L_{1j}, \dots , L_{6j}) \qquad \text{Eq. 2.3}$$

where $L_1, \dots , L_6$ are the educational categories of all males.[14]

---

[13]Because of the assumed perfect correlation between whites and nonwhites in corresponding educational categories, either the white or nonwhite educational categories could serve as well, but the labor types were defined in terms of all males and all females because of data limitations for the intended empirical test. Published data allow a distinction of race by the educational categories for relatively few

Finally, it can be useful to distinguish average earnings, e, according to sex:

$$e^{m}_{ij} = f(KL_j, L_{1j}, \ldots , L_{6j}) \qquad \text{Eq. 2.4}$$

$$e^{f}_{ij} = f(KL_j, L_{1j}, \ldots , L_{6j}) \qquad \text{Eq. 2.5}$$

where all arguments in both functions are as identified in Eq. 2.3 and the superscripts m and f signify male and female respectively. The hypothesis, stated in this form, allows for a comparison of the influence of cooperating factors on earnings by sex. If the females are in the same proportion to males for every educational category in every area as assumed, and if the theory is correct, the coefficients of the variables in Eq. 2.4 should be larger than and proportionate to coefficients of the corresponding variables in Eq. 2.5.[15]

---

metropolitan areas. Furthermore, in a slightly different formulation of the theory presented in this study several regressions were attempted which incorporated the distinction by race and all failed because of a high degree of multicollinearity.

[14]Alternatively, the educational categories could have been defined either in terms of the female labor force or the total labor force. Since it was considered desirable to allow for a sharp comparison of all three models and since Model III incorporates a distinction by sex, either the male or female categories would be preferable to the categories of the total labor force. Although either could serve as well empirically, the male categories were chosen simply because males generally comprise 2/3 of a metropolitan labor force.

[15]If the male-female ratio of employment is constant for corresponding educational categories in every area, the distinction by sex is completely valid without any qualification of the model. As judged by the data to be used laber, the ratios are near-constant, since the correlation coefficient varies from .836 to .963. Thus, proportionate coefficients are expected. Since the average earnings of males are higher than females for all labor types, the coefficients in the male equations should be larger.

A rationale for this model, as expressed in Eqs. 2.4 and 2.5, can be demonstrated by considering the general linear homogeneous production function

$$Q = F(K, L_1, \dots, L_6, L_f) \qquad \text{Eq. 2.6}$$

where Q = output; K = capital; and L = labor. The subscripts 1 to 6 denote the six educational categories of the male labor force, and the subscript f denotes female.

The marginal product of each factor input can be obtained by differentiating Eq. 2.6 successively with respect to each factor input:

$$MP_K = \partial Q/\partial K;\ MP_{L_1} = \partial Q/\partial L_1;\ \dots;\ MP_{L_6} = \partial Q/\partial L_6;$$

$$MP_f = \partial Q/\partial L_f. \qquad \text{Eqs. 2.7}$$

The function is linear homogeneous, and as a consequence Euler's theorem applies. If each factor is remunerated by an amount equal to its marginal product, the sum of the factor remunerations exactly exhausts the total output.

$$Q = (\partial Q/\partial K)K + (\partial Q/\partial L_1)L_1 + \dots + (\partial Q/\partial L_6)L_6 +$$

$$(\partial Q/\partial L_f)L_f \qquad \text{Eqs. 2.8}$$

where

$$r = \partial Q/\partial K;\quad w_1 = \partial Q/\partial L_1;\ \dots;\ w_6 = \partial Q/\partial L_6;\quad w_f = \partial Q/\partial L_f,$$

---

and r, $W_1, \ldots, w_6, w_f$, denote the returns to capital, the wage rate for each labor type, and the wage rate for women.

The relationship between productivity (and implicitly the wage rate) for each type of labor and both capital and the other types, including females, must still be made explicit. Since Eq. 2.6 is linear homogeneous, it can be written:

$$Q = Lf(K/L, L_1/L, \ldots, L_6/L, L_f/L) \qquad \text{Eq. 2.9}$$

where L = the total labor force. In this expression, output is a function of the capital-labor ratio, the male labor types, and the number of females expressed as a percentage of the labor force.

If the differences in wage rates are related to the ratios in this equation, the a priori expectation is that all relations are positive except for the type of labor being analyzed. For example, it is expected that $\partial e_{1j}/\partial(K/L) > 0$, that $\partial e_{1j}/\partial(L_i/L) > 0$ for all $i \neq 1$ and < 0 for $i = 1$. A similar expectation could be stated for all other labor types and the percent female. However, in a strict sense, the function does not allow for the differentiation of the educational categories or the percent female since the percentages are interdependent. They always sum to 1, so the percentage in any one category cannot be varied without simultaneously affecting the percentage in at least one other category or the percent female. Therefore, the expected sign can be stated only under the assumption that at least one category is allowed to vary along with the variable under consideration while the other variables remain constant. As a consequence, the expected sign will depend on the category chosen to vary along with the category under consideration.

This can be clarified in an analysis of a possible explicit relationship between Q and the other variables in Eq. 2.9:

$$Q = \beta_1(K/L) + \beta_2(L_1/L) + \ldots + \beta_7(L_6/L) + \beta_8(L_f/L) \qquad \text{Eq. 2.10}$$

It is reasonable to assume that all β's are positive, and if L_1 is identified as the lowest category (0-7) with each subscript identifying the next highest category, it is also reasonable to assume that $\beta_2 < \beta_3 < \beta_4 < \beta_5 < \beta_6 < \beta_7$.

Since $L_1/L + \ldots + L_6/L = PM$ (percent male), Eq. 2.10 can be written

$$Q = \beta_1(K/L) + \beta_2 PM + (\beta_3-\beta_2)L_2/L + \ldots + (\beta_7-\beta_2)L_6/L + (\beta_8-\beta_2)L_f/L. \qquad \text{Eq. 2.11}$$

With the equation in this form it is clear that the sign of any educational category depends upon the category which is allowed to vary. In this equation, L_1 is allowed to vary and therefore the signs of all the remaining categories can be expected to be positive. If L_6 were the category allowed to vary, the signs of the remaining categories could be expected to be negative.

The sign of L_f/L will also depend on the value of β_8 relative to the value of the category which is allowed to vary. This value will depend upon the educational composition of the female labor force relative to the male labor force. If the educational composition of the female labor force is radically diverse and random from area to area, nothing definite can be stated about the value of β_8. The sign of L_f/L must then be predicted as contingent, i.e., as negative if $\beta_8 < \beta_2$, as positive if $\beta_8 > \beta_2$, and as zero if $\beta_8 = \beta_2$.

A definite sign can be affirmed with different assumptions. On the basis of the traditional assumption of woman as the weaker sex and in light of previous discussion, an assumption can be made that women in corresponding categories are less productive than men.[16] Given this assumption and the assumption that all categories vary proportionately with any change in L_f/L, i.e., if a change in L_f/L requires a change in the male-female ratio of every educational category such that the percentage of the total labor force in each category remains constant, the expected sign is negative. But it is to be noted that the sign is affirmed without establishing the value of β_8 relative to the coefficients of the categories. This is unnecessary because of the assumption that all categories vary with a change in L_f/L. Such an assumption can be supported for an empirical test with a high correlation

[16]This assumption is certainly not advanced as valid, but is used to create an analytical mode to test, in part, propositions concerning female as opposed to male wage determinants.

between corresponding categories.

This discussion of an expected sign for L_f/L was carried on in anticipation of Model III. The variable is assumed constant in Model I and was therefore not included in Eqs. 2.4 and 2.5. The previous considerations define the signs of the KL_{ij} and L_{ij} variables in those two equations.

In summary, Model I, as expressed in Eqs. 2.4 and 2.5, is built on the assumption that the male-female and white-nonwhite ratios are identical for every area; that the influence of unionization is identical in each area; and that the percentages in each educational category express equilibrium supply in each area as determined by the demand for each labor type. The latter assumption will be relaxed in Model II and the prior assumption will be relaxed in Model III. As a result of the assumptions for Model I, each area can be considered in equilibrium internally, with different factor ratios as the only source of earnings differentials between areas.

2. Model II

Job requirements for any given industry are generally set from an analysis of experience with laborers who have performed at the same or similar tasks. Hiring criteria, including educational level, are determined by this analysis. Thus, the educational composition of an industry is rooted in its technology, and each industry can be identified by the educational compositon of its labor force.[17] It can therefore be affirmed that the educational composition of an area's labor force follows from the industrial structure. On the basis of this affirmation, this structure can serve to identify an expected or required labor force educational composition.

It was assumed in Model I that the labor types are remunerated by an amount equal to their respective marginal products. Implicit in this assumption is the identification of the demand function for a labor type with its marginal product function. Consequently, if the supply of a labor type is less than an equilibrium quantity relative to the demand as determined by the industrial structure, the marginal product and therefore the earnings of that labor type will be higher than they would be at a quantity consonant with the structure. The reasoning is symmetrical for a labor type whose quantity is greater than that required by the structure. In both cases Model I will suffice if the marginal product function completely specifies demand for a labor type regardless of the area industry requirements. However, if a labor type is important to the industrial process and cannot be substituted for easily by other labor types, it is possible

---

[17] 1960 Census data show that the labor forces of different industries are different in educational composition, and the identification of an industry with the educational composition of its labor force is supported in the research of Victor R. Fuchs: "... in a study of interindustry wage differentials now under way at the National Bureau, we find that the terms "high paying" and "low paying" can be profitably replaced by identifying those characteristics such as color, age, sex, and education of the labor force, location, extent of unionization, and size of employer that explain nearly all of the interindustry differences in earnings." Victor R. Fuchs, op. cit., p. 2.

that area firms will pay more in the short run than is indicated by the marginal product function to secure the requisite skills in order to continue producing with higher efficiency. This possibility seems more likely with highly technical or highly skilled jobs, but a variable can be added to the equations ror every labor type in Model I to indicate an over- or under-supply relative to the industrial structure. Under the assumptions of importance to the industrial structure and limited substitution possibilities, it is expected that an excess supply of labor type will decrease its wages and an under-supply will increase its wages. A formal specification of the demand function is unnecessary since a variable can be introduced into the analysis which can capture the effects of the adjustment process of a short-run disequilibrium without such a specification. This could be done by a variable which would indicate whether the relative percentage of an educational category in an area is less than, equal to, or more than an equilibrium quantity as determined by the industrial structure. Such a variable can be added to each equation of Model I, and for any particular labor type it can be positive in sign, equal to zero, or negative. Since it can be viewed as measuring excess or deficient demand for the labor type, it can be referred to as a demand variable. It should capture a variation in earnings not accounted for by relative factor supplies.

Model II can be stated formally:

$$e^m_{ij} = f(KL_j, L_{1j}, \ldots, L_{6j}, D^m_{ij}) \qquad \text{Eq. 2.12}$$

$$e^f_{ij} = f(KL_j, L_{1j}, \ldots, L_{6j}, D^f_{ij}) \qquad \text{Eq. 2.13}$$

where the demand variable, D, is defined as the difference between the equilibrium supply of a labor type and the actual supply. It is conceived as being calculated by subtracting the actual percentage of the labor type from what can be considered the equilibrium percentage as determined by the industrial structure. Given the assumptions and expectations above, the expected sign of D is positive; i.e., $\partial e_{ij}/\partial D_{ij} > 0$.

This model differs from Model I only in the assumption that the relative percentages of the educational categories do not necessarily express equilibrium supply of each labor type relative to demand. The adjustment process due to over- or under-supply is therefore another possible source of inter-area differentials.

3. Model III

Factor ratios are a basic source of interarea earnings differentials in both previous models. In the first model, differences in factor ratios between areas provide the only explanation for the differential. In the second model, another possible source is the difference between the equilibrium supply and the actual supply of a factor within an area. The third model adds three other possible sources: differences in the male-female ratio of the area labor force; differences in the white-nonwhite ratio of the area labor force; and differences in the degree of unionization between areas.

It was reasoned above that both females and nonwhites could be considered as fundamentally different resources from males and whites. On this basis the labor types in Eq. 2.1 were distinguished according to both sex and race. However, both previous models were developed on the assumption of identical white-nonwhite ratios for both males and females and identical male-female ratios for every educational category in every area. The former assumption was made because an analysis of the data for the empirical portion of this study revealed that too few metropolitan areas have enough nonwhites in the labor force to make a test of a model based on Eq. 2.1. The latter assumption was made because an analysis of the same data revealed that the correlation between corresponding male and female educational categories was very high.[18]

But neither of these findings demands the conclusion that the earnings of the individual labor types are unrelated to the proportion of all females or nonwhites in the labor force. Both the percent female and the percent nonwhite can be incorporated into the analysis in either of two ways.[19]

First, the assumption of the previous models could be changed to allow L_f/L in Eq. 2.9 to vary and, without specifying the structural relationship, an a priori assumption could be made that the earnings of all labor types are related to the percent nonwhite in an area with the nature of the relationship being determined by other assumptions.

With L_f/L being allowed to vary, its sign can be predicted on the basis of the discussion carried out previously; i.e., as contingent if the educational composition of the female labor force is diverse and random from area to area, or as definitely negative if women are assumed less productive than men and if all categories vary proportionately as to male-female composition with any change in L_f/L. The high correlation between corresponding male and female categories in the data for the empirical test supports the expectation of a negative sign as more appropriate.

---

[18]In a slightly different formulation of the theory presented in this study, several regressions were attempted which incorporated the distinction by sex, and all failed because of a high degree of multicollinearity.

[19]A high correlation between corresponding male-female categories does not force the inference that the percentage of females or nonwhites in the area labor force is empirically meaningless.

On the assumption that the educational composition of both male and female nonwhites is proportionate to the white male and female categories, the earnings of all labor types could be assumed negatively related to the percent nonwhite. This is predicated on the supposition that the nonwhite is a "weaker" labor type (See p. 7). Under these assumptions a change in the percent nonwhite is reflected in a relative change in the white-nonwhite ratio in every educational category, but the relative percentages of laborers in each category of the male and female labor forces remains the same.

Even if these assumptions are not perfectly true, it is doubtful whether the actual differences in educational composition of the nonwhite labor force between areas are significant.

In the period to be studied, nonwhites in fact were employed primarily in the lower-paid occupations. As Herman P. Miller found in his study of 1960 Census data,

> A nonwhite man who has not gone beyond the eighth grade has very little chance of being anything more than a laborer, porter, or factory hand. ... The nonwhite high school graduate stands a somewhat better chance of getting a well-paid job; but even his chances are not very good. About 6 out of every 10 nonwhite high school graduates were laborers, service workers, or operatives ... relatively few nonwhites were in the higher paid professions ... only 8 percent.[20]

---

[20]Herman P. Miller, Income Distribution in the United States, (1960 Census Monograph), U.S. Government Printing Office, 1966, pp. 163-164. It might be mentioned that there is indication from more current data that this is no longer true. See Duran Bell, "Occupational Discrimination as a Source of Income Differences: Lessons of the 1960's," American Economic Review, 62, May 1972, p. 363.

This indicates that nonwhites in fact were generally treated as a weaker resource, and it supports the contention that, other things equal, a larger percentage of nonwhites in an area will result in lower labor type earnings.

Both the percent female and the percent nonwhite can enter the analysis in an alternate but equally valid way. Equation 2.9 can be disaggregated to distinguish the percent nonwhite:

$$Q = Lf(K/L, L_1/L, \ldots , L_6/L, L_f/L, L_n/L) \quad \text{Eq. 2.14}$$

where $L_1/L, \ldots , L_6/L$ are now defined as <u>white</u> male labor types, L_f/L as the percent of <u>white</u> females, and L_n/L as the percent of nonwhites.[21] The expected signs are the same as for the corresponding variables in Eq. 2.9. On the basis of the assumptions and the analysis carried out above, the expected sign of L_n/L is negative. If these assumptions are changed so that the educational composition of the nonwhite labor force is diverse and random from area to area, the sign of L_n/L is contingent upon the productivity of whites relative to nonwhites in corresponding educational categories and on the coefficient of the category allowed to vary. This can be made explicit by rewriting Eq. 2.14 in a form comparable to Eq. 2.11.

---

[21]As stated previously, data limitations do not allow the model to be tested in this form. However, the high correlation of these variables with the corresponding categories of the male labor force, etc., which are used in Models I and II, justifies the use of the categories described in those models for an empirical test.

$$Q = \beta_1(K/L) + \beta_2 PM + (\beta_3-\beta_2)L_2/L + \ldots + (\beta_7-\beta_2)L_6/L$$

$$+ (\beta_8-\beta_2)L_f/L + (\beta_9-\beta_2)L_n/L. \qquad \text{Eq. 2.15}$$

In this equation it is clear that the sign of L_n/L is dependent upon the coefficient of L_n/L as could be identified in an explicit statement of Eq. 2.14 and the coefficient of the category which is allowed to vary. But on the basis of the reasoning accompanying the assumptions for the first approach to incorporating L_n/L into the model, the expected sign will be stated as definitely negative.

As discussed above, differences in the degree of unionization may be important in explaining area wage differentials. If two areas are identical in resource structure and they experience the same relative demand for the various labor types, it is expected that all labor types in the more highly unionized area will have higher earnings than the labor types in the less unionized area.

A formal statement of Model III is identical to Model II, as expressed in Eqs. 2.12 and 2.13, except for the addition of a sex, race, and union variable:

$$e^m_{ij} = f(KL_j, L_{1j}, \ldots, L_{6j}, D^m_{ij}, PF_j, PN_j, PU_j) \qquad \text{Eq. 2.16}$$

$$e^f_{ij} = f(KL_j, L_{1j}, \ldots, L_{6j}, D^f_{ij}, PF_j, PN_j, PU_j) \qquad \text{Eq. 2.17}$$

where PF = percent female in the labor force, PN = percent nonwhite in the labor force, and PU = percent of the labor force under union contract. The signs of these variables are as indicated in the previous discussion; i.e., $\partial e_{ij}/\partial PF_j < 0$, $\partial e_{ij}/\partial PN_j < 0$, and $\partial e_{ij}/\partial PU_j > 0$.

4. Regional Differences

A modification of all three models is inspired by the differential in money earnings between northern and southern workers which was discussed in the introduction to this study. This differential is one of the more enduring phenomena ever examined by economists. It has been the subject of analysis for over forty years and there is reason to believe that the analysis will continue for some time.

In the original study, Clarence Heer concentrated on the period from 1849 to 1927 and found that the average wage in the South for textile workers was 40 percent below that obtained in other parts of the United States.[22] Joseph W. Bloch, in a regional analysis made just subsequent to the Second World War, concluded that in 1907 northern workers were earning twice as much as southern workers for similar tasks.[23] Victor R. Fuchs, analyzing data from the 1960 Census of Population, found that average hourly earnings for all non-agricultural

---

[22]Clarence Heer, Income and Wages in the South, University of North Carolina Press, Chapel Hill, North Carolina, 1930.

[23]Joseph W. Bloch, "Regional Wage Differentials, 1907-1946," Monthly Labor Review, April 1948, 66, pp. 371-377.

workers, including the self-employed, were about 20 percent lower in the South than in the rest of the nation.[24] In a similar study, Harry M. Douty concluded that there was relatively little change in the differential by 1963.[25]

Most recently, two studies have claimed a complete or near-complete explanation of the differential. Gerald W. Scully analyzed industry wage variation between states and was able to explain approximately 90 percent of the differential with five variables: industry capital-labor ratio, education of industry labor force, percent female in industry labor force, perent nonwhite in industry labor force, and percent of industry labor force under union influence.[26]

As did all previous researchers, Scully measured the differential by money earnings. Philip R. P. Coelho and Moheb a. Ghali deflated the data for five northern and five southern metropolitan areas and found no difference in earnings.[27] The differential indicated by money earnings disappeared. This suggested that real earnings are identical in all of the North and all of the South.

---

[24]Victor R. Fuchs, Differentials in Hourly Earnings by Region and City Size, 1959, National Bureau of Economic Research Occasional Paper 101, Columbia University Press, New York, 1967.

[25]Harry M. Douty, "Wage Differentials: Forces and Counterforces," Monthly Labor Review, March 1968, 91, pp. 74-81.

[26]Gerald W. Scully, "Interstate Wage Differentials: A Cross-Section Analysis," American Economic Review, 59, December 1969, pp. 757-773.

Coelho and Ghali's study was challenged by Mark L. Ladenson, primarily on the basis of their choice of southern metropolitan areas. Ladenson replaced four of their five southern cities with cities that "would be recognized as southern by any reasonable observer," and the North-South differential reappeared in the deflated data: 13.1 percent in hourly real wages and 10.6 percent in annual real wages.[28]

Coelho and Ghali claimed that Ladenson's conclusions are faulty because of his arbitrary exclusion of available data, his choice of a low budget cost of living as a deflator, and the possibility of heteroscedasticity which would render his testing procedure invalid.[29]

This reply to Ladenson presents the most recent addition to the voluminous literature on the North-South differential. Further studies will very likely follow when deflators are available for a larger number of cities, for the small number of observations encourages great caution in the acceptance of Coelho and Ghali's conclusions.

On the basis of these non-conclusive results and the long history of the differential, a dummy variable will be

---

[27]Philip R. P. Coelho and Moheb A. Ghali, "The End of the North-South Wage Differential," American Economic Review, December 1971, 61, pp. 932-937.

[28]Mark L. Ladenson, "The End of the North-South Wage Differential: Comment," American Economic Review, 63, September 1973, pp. 754-756.

[29]Philip R. P. Coelho and Moheb A. Ghali, "The End of the North-South Wage Differential: Reply," American Economic Review, 63, September 1973, pp. 757-762.

added to each of the previous models to test whether there is an earnings differential between northern and southern cities for any of the six labor types which cannot be explained by one or more of the models.

CHAPTER THREE

EMPIRICAL TEST OF THE MODELS

An empirical test of the hypotheses embodied in the several models requires data that will yield the following: 1) average money earnings of labor types in each area identified by years of education;[1] 2) a measure of the total physical capital in each area; 3) a measure of the total area labor force; 4) percentages of each male labor type in the area labor force as identified by years of education; 5) the percentage of females in the area labor force; 6) the percentage of nonwhites in the area labor force; and 7) the degree of unionization in each area. The 1960 *Census of Population*, the 1967 *Census of Manufactures*, the 1967 *County and City Data Book*, and the Bureau of Labor Statistics

---

[1]Money earnings will be used primarily because of a lack of deflators for the areas to be studied. However, there is good reason to accept the proposition that relative real earnings are insignificantly different from relative money earnings for 1959, the year to be studied. In a consideration of the same problem for his study of earnings differentials, which concentrated on the year 1959 also, Victor R. Fuchs found that "fragmentary information provided by the Bureau of Labor Statistics on the cost-of-living in different cities suggests some slight correlation between hourly earnings and prices, but intercity differences in cost-of-living appear to be small relative to differences in hourly earnings." See Victor R. Fuchs, *op. cit.*, p. 34. This finding is consistent with the reflections of Borts and Stein: "... a wage differential of x percent between regions could not produce a difference in cost-of-living of x percent for two reasons: first, many items in the consumer budget are produced outside the low-wage region and consequently are influenced in price by the wage levels of other regions; second, many items in the consumer budget are produced by capital cooperating with labor. If capital is not receiving a lower rate of return in the low-wage region, the relative cost of living in the two regions will not be as low as the relative wages. Because of these considerations, differences in the cost of living can never completely offset the influence of wage differences." See George H. Borts and Jerome L. Stein, *Economic Growth in a Free Market*, Columbia University Press, New York, 1964, p. 11

Bulletins of Wages and Related Benefits contain data sufficient for the test. Forty-five (45) Standard Metropolitan Statistical Areas (SMSA's) over 250,000 population, as defined by the Bureau of the Census for 1959, are the areas to be studied.[2]

1. Data for Average Earnings of Labor Types

The average earnings of the labor types in each area can be estimated from the median annual earnings in 1959 for detailed occupations in each SMSA, published in the 1960 Census of Population, and the national percentages in each educational category for every occupation in 1959, which can be calculated from data published in Subject Report PC (2)-7A, Occupational Characteristics, derived from the 1960 Census of Population. An estimation of some type is necessary since the exact earnings of the labor types are not published and they cannot be calculated from published data. The following will be used as the formula of estimation:

$$e_{ij} = \sum_m n_{ijm} w_{jm} \Big/ \sum_m n_{ijm} \qquad \text{Eq. 3.1}$$

where

n_i = the number of workers in the occupation in the SMSA with educational level i. This is estimated by multiplying the total number of workers in the occupation in the SMSA by the national proportion of the workers in the occupation with educational level i in 1959.

w = median annual earnings of the occupation in the SMSA in 1959

[2]The number of observations is limited by the data for both the capital-labor ratio and the degree of unionization. Data for 1959 are

j = SMSA

m = the number of occupations in the SMSA in 1959.

This variable may be somewhat misspecified because of the national weights, but there is no reason to assume any systematic bias. As indicated in the several models, separate variables will be computed for males and females. The calculations will involve as many as 159 detailed occupations for males and as many as 68 detailed occupations for females in each SMSA.

2. Data for the Capital-Labor Ratio

Data for the amount of physical capital by SMSA for 1959 is not available, but data is available which will allow the computation of capital-labor ratios in manufacturing as proxy variables for the SMSA capital-labor ratios.

One possible proxy is the amount of value added by SMSA, published for 1958 in both the Census of Manufactures and the 1967 County and City Data Book. Two assumptions are necessary to make this an acceptable proxy: 1) that the amount of value added in 1958 is not substantially different from that in 1959 for each SMSA, and 2) that the amount of value added is proportional to the capital stock. Both assumptions are weak. The year 1958 was a business cycle trough and value added is a weak proxy for the capital stock because it includes the effect of the skill level and union power as well.

---

used because the data collection was completed prior to the publication of 1970 Census data, and an analysis utilizing the earlier data has both its own validity and can provide a basis for later comparison.

as the amount of available capital per worker.[3]

Another possible proxy is the amount of new investment expenditures, averaged over several years. For this to be an acceptable proxy it is necessary to assume that the average amount of new investment in an SMSA over several years is proportional to the capital stock of the area.[4] This is not an unreasonable assumption if the influence of the business cycle can be offset. Replacement investment should be a relatively constant proportion of the total capital stock and it should react rather slowly to the cycle, whereas net investment should be the more volatile component of gross investment.

---

[3]Value added was used as a proxy for the capital stock in S. H. Masters, "Wages and Plant Size: An Interindustry Analysis," Review of Economics and Statistics, 51, August 1969, and the author was criticized for the reasons listed. For the criticism, see C. T. Haworth and D. W. Rasmussen, "Human Capital and Interindustry Wages in Manufacturing," Review of Economics and Statistics, 53, November 1971, pp. 376-379.

[4]Lawrence R. Klein discusses some of the difficulties in any attempt to grasp the capital stock with existing data: "The aggregation of capital would be like the aggregation of other economic variables were it not that different vintages of capital goods have different technical attributes. ... In national income statistics, we are usually furnished with periodic estimates of gross investment in fixed capital and capital consumption or depreciation. The latter are usually based on book value accounting estimates and not actuarial estimates. Gross investment, less depreciation, gives net investment, and successively cumulated totals of the latter from some original asset levels give series of capital stock. ... To measure real capital from this formula, all the component variables must be expressed in a constant price system. We are still, however, faced with the problem of quality change and the fact that all the different vintages combined in this formula are of heterogeneous quality." See Lawrence R. Klein, Introduction to Econometrics, Prentice-Hall, New Jersey, 1962, pp. 87-88. The measure suggested here -- average amount of investment in an SMSA over several years -- is not identical to that discussed by Klein, but it has as many weaknesses. However, we know of no better measure available.

Data is not available for every year and every SMSA, but the national figures for capital consumption allowances (the demand for replacement investment) shows a high degree of constancy and gross investment reflects the business cycle trough (1958) and then shows an erratic increase in boom years when measured as a percent of GNP. See Table 3.1.

TABLE 3.1

Percentage of Gross Income

Year	Capital Consumption Allowance	Gross Private Domestic Investment
1954	7.7	14.2
1955	7.9	16.9
1956	8.1	16.7
1957	8.4	15.4
1958	8.7	13.6
1959	8.6	15.6
1960	8.6	14.8
1961	8.7	13.8
1962	8.9	14.8
1963	9.2	14.8
1964	8.9	14.9
1965	8.7	15.8
1966	8.5	16.2

Source: The National Income and Product Accounts of the United States, 1929-65, Statistical Tables, U.S. Department of Commerce, 1966, and Survey of Current Business, July 1969, U.S. Department of Commerce

Admittedly, the capital consumption allowance is an accounting estimate of actual depreciation, but the problem of measuring depreciation is old and only one of many in any attempt to measure the capital stock.[5] Although no one has

[5]Reflecting on this problem, Simon Kuznets states, "... all concepts in the field of national income are, in one way or another, non-operational.

specified the actual time shape of depreciation, it does appear to be a relatively constant proportion of GNP for the years which are of interest here. A proxy is needed for the capital stock of selected SMSA's for 1959, and the 1967 Census of Manufactures lists new investment expenditures for the years 1954, 1958, and 1963. The immediate impulse is to use the average for all three years, but it could be argued that an average of two years, one representing a business cycle trough and the other a business cycle boom, would give a better indication of the actual values. If this argument is accepted, an average of 1958 and 1963 would give the best index from the available data. In fact, a comparison of the three- and two-year averages reveals that there is no substantial difference between them. The two-year average will be used, however, since it yields more observations.

The same Census of Manufactures also lists the man-hours of production workers for 1959. Therefore the capital-labor ratio in manufacturing can be computed for each SMSA, a ratio of the average new investment expenditures for 1958 and 1963 to the total man-hours of production workers in 1959. This will serve as the proxy for the total SMSA capital-labor ratio.

They are goals that forever elude measurement and for which measurable approximations are substituted. ... All these operational measures assume meaning only because they are approximations of the 'purer' non-operational concepts behind them." See Simon Kuznets, "Comment" on Edward F. Denison, "Theoretical Aspects of Quality Change, Capital Consumption, and Net Capital Formation," Problems of Capital Formation, National Bureau of Economic Research, New York, 1969, p. 72.

3. Data for the Educational Categories, Percent Female, and Percent Nonwhite

A measure of the level of education of the area labor force is necessary to test the hypotheses as stated, but such data has not been published. Information on the education of the labor force by SMSA for 1959 is not available. However, the 1960 Census of Population does list the level of education for all males and females over 25 in each SMSA with a population over 250,000. This data can be used to test the hypotheses if it can reasonably be assumed that the educational configuration of this group is highly similar to that of the labor force.

An examination of data from twenty SMSA's, chosen randomly, revealed that the labor force in 1959 was comprised of from 35 to 40 percent of the area population and that males outnumbered females in a ratio of 2 to 1. Also, a large portion of the area labor force was under 25 years of age.

Because of the greater proportion of the population completing both high school and college since the Second World War, it is possible that there is a slight downward bias, especially in the higher educational categories, when using only the over 25 group as a proxy for the total labor force. However, it is probably not enough to make a substantial statistical difference, particularly if all SMSA's experience a similar bias.

A more important problem arises with a possible sex bias. An unqualified analysis of the area population would give women too great a weight in any percentage computation, since women comprise approximately 1/2 of the population but only 1/3 of the labor force. Because of these relative ratios, the male categories are likely to better represent the labor force educational mix, and this was the basis for using them in both male and female equations of all three models.

An even more practical decision had to be made in choosing the base for the computations of the categories -- should they be calculated as a percent of the male population or as a percent of the total population over 25. Since the correlation coefficient between corresponding categories varies from .980 to .997, there is no substantive difference between the two ways of calculating the percentages. Either measure will yield the same information. But the choice of one in preference to the other is not completely without consequence.

The expected signs for the theoretical models were predicated on the assumption that one category would be allowed to vary together with the category under consideration, because for any n random categories which sum to the same total, the nth category is dependent on the values of the other categories. Therefore, in a statistical test, one category is omitted. Conceptually, it is the category which varies together with the category under consideration.

If the computations are made on the basis of the male population, the sum of the separate percentages for the six categories will always equal 100% for each area. As a consequence, one educational category must always be omitted. This same conclusion does not follow if the calculations are made in terms of the total population. With this base the percent female can be considered the omitted category and the six male educational categories can be considered independent.

However, the latter solution is not clearly superior. It necessarily excludes a direct measure of the influence of sex. The alternative is to omit one of the male categories and include the percent female, but this is equivalent to using five categories from the male population and separately calculating the percentage of the labor force which is female.

Since a direct measure of the influence of sex is more desirable than the influence of an extra educational category, and since the percent female is incorporated in Model III, either solution which allows for such a measure can be considered superior to the other alternative. At any rate, it is impossible to test the hypothesis as stated in Eqs. 2.16 and 2.17; i.e., with six educational categories and the percent female. Therefore it will be tested as suggested in the theoretical discussion of the models -- with five educational categories, calculated from the male population over 25 in each SMSA and the percentage of the labor force which is female.[6]

Both the percent female and the percent nonwhite of the area labor force can be calculated from data published in the 1960 Census of Population. It is assumed that percentages for April, 1960, are representative of 1959.

4. Data for a Measure of Area Unionization

Data for the extent of unionization by SMSA in 1959 is not available, so a proxy variable is necessary. An examination of past research has shown that two could be used in this study.

The first is the number of workers involved in work stoppages computed as a percent of the area labor force. Gerald W. Scully used this in one of his research projects, and he explains both its rationale and its weaknesses:

> Union activity, expressed as the percentage of workers involved in work stoppages which is the specification of the union variable used in this study, may vary from year to year for a variety of reasons. Collective bargaining agreements are often concluded for a period longer than a year, and the variable may shift over the trade cycle. Thus, economic factors in any given year may be such that the number of workers involved in work stoppages may be unusually large or small. To obviate this difficulty, the mean value of the workers involved in work stoppages over the period 1956-60 is employed. Until other data becomes available

---

[6]Since the correlation between corresponding categories is almost perfect (.980 to .997), the coefficients of the five educational categories calculated as percentages of the total labor force could be easily derivable, with slight error, from the coefficients of the categories calculated as percentages of the male labor force. However, the absolute value of a coefficient will not be as interesting as its size relative to the other categories, and this will be almost identical regardless of which percentages are used.

> this variable would seem to be an adequate measure of the influence of union activity on interstate wage differentials. It should be noted that it has its weaknesses. First, it may measure more than pressure to bid up wage rates: work stoppages can result from unionization membership drives and jurisdictional disputes. Second, a higher percentage of workers involved in work stoppages may occur in low wage paying states reflecting an effort to bring wage rates into parity with other regions. While these two limitations are recognized, it is felt that higher percentages of workers involved in work stoppages will reflect the operation of trade union market forces and will be positively associated with higher wage rates.[7]

The data are readily accessible in one publication, BLS Report No. 236, Work Stoppages for Metropolitan Areas, 1952-62, and the percentages can be easily computed by dividing the total area labor force into the number of workers involved in a work stoppage.

An alternative proxy is the percentage of an area's labor force covered by union contract for 1960 and later years as published in the Bureau of Labor Statistics' Wages and Related Benefits, Bulletins 1285-83, 1303-83, 1430-83 and 1575-87. Since the percentages have remained fairly constant for 1960 and the several following years, it could be assumed that the percentages were also relatively the same for 1959.

This second proxy has a weakness, but this weakness does not seem too important. The percentages of an area's labor force covered by union contract are not listed with precision; they are expressed in terms of five-percentage-

[7]Gerald W. Scully, "Interstate Wage Differentials: A Cross-Section Analysis," American Economic Review, 59, December 1969, p. 760.

point ranges. For example, the percent of plant workers employed in Boston who were covered by union contract is listed as 70-74 percent. Similar ranges are given for other areas. Thus, the actual difference between the unionization of two SMSA's may be absorbed in a procedure such as selecting the middle value of the range. But in comparing the strengths and weaknesses of the two proxies, this second proxy seems much stronger and will be used to measure the extent of SMSA unionization.

5. Data for the Demand Variable

Model I assumes that the relative percentages of each educational category express equilibrium supply of each labor type. If they do not, which is the assumption of both Model II and III, excess demand or supply can be indicated by the difference in the actual percentages relative to an index of equilibrium size. Since each industry requires a certain educational composition of its labor force, as discussed previously, the industrial structure of an area can serve as a basis for deriving an expected labor force composition. Therefore, the following variable should serve adequately as an index of demand for each labor type:

$$D_{ij} = (\sum_{k} N_{ijk} / \sum_{i,k} N_{ijk}) - L_{ij} \qquad \text{Eq. 3.2}$$

where

D_i = demand for labor type $\underline{i}$.

N_{ij} = the number of workers in the industry that should have achieved educational level i in the jth SMSA. This is estimated by multiplying the total number of workers in the industry by the national proportion of that educational level in the industry, which can be calculated from 1960 Census data.

k = the number of industries in the SMSA in 1959.

L_{ij} = the actual percentage of the labor force (male or female, depending on the dependent variable) which has achieved educational level i in the jth SMSA.

These demand variables will be positive, negative, or zero depending on whether the proportion of the labor force that is expected to have achieved a given educational level is greater than, equal to, or less than the actual proportion. The calculations will involve as many as 147 industries for the male labor force in each SMSA and as many as 122 for the female labor force.

6. The Empirical Models

The previous discussion provides a basis for testing the hypotheses expressed in the several models. The equations as modified by subsequent discussion suggest least-squares regressions of the following form:

Model I

$$e^m_{ij} = a + bKL_j + c_1L_{1j} + \ldots + c_5L_{5j} + u. \qquad \text{Eq. 3.3}$$

$$e^f_{ij} = a + bKL_j + c_1L_{1j} + \ldots + c_5L_{5j} + u. \qquad \text{Eq. 3.4}$$

Model II

$$e^{m}_{ij} = a + bKL_{j} + c_{1}L_{1j} + \ldots + c_{5}L_{5j} + dD^{m}_{ij} + u. \quad \text{Eq. 3.5}$$

$$e^{f}_{ij} = a + bKL_{j} + c_{1}L_{1j} + \ldots + c_{5}L_{5j} + dD^{f}_{ij} + u. \quad \text{Eq. 3.6}$$

Model III

$$e^{m}_{ij} = a + bKL_{j} + c_{1}L_{1j} + \ldots + c_{5}L_{5j} + dD^{m}_{ij} + ePF_{j} + fPN_{j} + gPU_{j} + u. \quad \text{Eq. 3.7}$$

$$e^{f}_{ij} = a + bKL_{j} + c_{1}L_{1j} + \ldots + c_{5}L_{5j} + dD^{f}_{ij} + ePF_{j} + fPN_{j} + gPU_{j} + u. \quad \text{Eq. 3.8}$$

where

e = average annual earnings of a labor type identified by years of education and computed according to the formula expressed in Eq. 3.1.

KL = the ratio of the average new investment for the years 1958 and 1963 to manhours of labor for 1959.

D = a demand variable calculated according to the formula expressed in Eq. 3.2.

L = the educational categories of the male population over 25 calculated as a percentage of the same population in 1959.

PF = percent female of the labor force in 1960.

PN = percent nonwhite of the labor force in 1960.

PU = percent of the labor force covered by union contract calculated as an average of 1960, 1961, and 1962.

The superscripts m and f signify male and female respectively. The subscripts i and j signify labor type identified by educational category and SMSA respectively.

As indicated previously, a second set of regression equations will be calculated with the addition of a regional dummy variable, North = 0, South = 1.

According to the theory, the signs of KL, PU, and D are expected to be positive; the signs of L are all expected to be positive when L_1 is the omitted category and negative when L_6 is the omitted category; and the signs of PF and PN are expected to be negative.

CHAPTER FOUR

THE EMPIRICAL EVIDENCE

The results of estimating the regressions for the three models as expressed in Eqs. 3.3, ... , 3.8 are contained in Tables 4.1, 4.2, and 4.3. For the purpose of easy comparison, the results are separated by sex within each model.

Before attempting an interpretation, it should be noted that variables were not added or deleted to obtain better statistical results -- a set number of variables was included in every equation according to the theory previously specified. Such a procedure may result in problems of multicollinearity. Admittedly, simple correlation is not an unfailing index of multicollinearity, but the correlation matrix of the various independent variables suggests that some multicollinearity may be present which would bias the coefficients.[1]

In all subsequent discussion, a variable will be referred to as "significant" if the test supports that statement at the five-percent or at a higher level. If a different level of significance is to be considered, a qualifying statement will be made.

---

[1]Rao and Miller state that: "A standard rule that some investigators have been using calls for inspection of the simple correlations among the independent variables. One should realize that simple correlations are only elements of the entire correlation matrix and, hence, may or may not contribute to problems of multicollinearity. One should not, a priori, rule out estimation of regression equations because of high simple correlations between any two independent variables." Potluri Rao and LeRoy Miller, Applied Econometrics, Wadsworth Publishers, Belmont, Calif., p. 48. The correlation coefficients were generally below .40, but, e.g., the coefficient for PN and the 0-7 educational category is .71.

Another clarification may be helpful. Two variables are related to a portion of the labor force described by level of education: average earnings (the dependent variable in the regressions) and the percentage of the male labor force having completed a given level of education (independent variables in the regressions). In order to avoid confusion, the term "labor type" will be used when referring to the labor group identified by the dependent variable and the term "educational category" will be used when referring to the educational percentages which define the independent variables.

1. An Overview

As indicated by the F-test, all equations for every model are significant at the one-percent or at a higher level.[2] Most notably, the theory is not completely supported in any one regression equation. The signs of the variables are not always consistent with the theory, and in several instances the signs change from model to model. Only the 0-7 educational category and PU have both the correct sign and possess a high degree of significance in all equations for all models.[3] All other variables are inconsistent in sign for at least one equation in one model and the significance of some changes from model to model.

---

[2]As calculated with the different degrees of freedom for each regression, the critical value for F varies from a low of 2.9 to a high of 3.3 in a test at the one-percent level. The actual F-values for the regressions vary from a low of 4.78 to a high of 21.6 with almost all having a value of 7.0 or higher.

TABLE 4.1

MODEL I REGRESSION EQUATIONS FOR MALES

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	D	$\bar{R}^2$
0-7M	8325.31	-0.22 (-0.08)	-102.88 (-3.57)*	5.32 (0.18)	27.12 (0.69)	-77.44 (-1.98)*	-69.22 (-1.32)			.77
8M	-1130.13	1.44 (0.54)		99.38 (5.79)*	112.95 (3.92)*	18.46 (0.83)	46.21 (1.40)	99.55 (3.64)*		.76
9-11M	-626.23	1.37 (0.53)		92.18 (5.34)*	103.11 (3.69)*	14.81 (0.69)	42.66 (1.33)	103.47 (3.90)*		.75
12M	444.14	1.73 (0.67)		77.47 (4.63)*	87.54 (3.11)*	8.33 (0.38)	46.46 (1.44)	109.17 (4.09)*		.71
13-15M	1500.67	1.17 (0.41)		62.24 (3.40)*	79.74 (2.59)*	-2.77 (-0.12)	32.52 (0.92)	133.38 (4.57)*		.66
16M	3335.83	1.06 (0.39)		42.25 (2.41)*	76.34 (2.59)*	-10.99 (-0.49)	-3.57 (-0.11)	154.99 (5.55)*		.67

t-values are shown below coefficients (critical value = 1.69; 38 d.f.)

* significant at the 5% or at a higher level with the correct sign

● significant at the 5% or at a higher level with the wrong sign

TABLE 4.

MODEL I REGRESSION EQUATIONS FOR FEMALES

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	D	$\bar{R}^2$
0-7F	5596.93	-3.78 (-2.29)●	-71.59 (-4.22)*	-10.67 (-0.60)	14.19 (0.61)	-78.23 (-3.39)*	-53.81 (-1.74)*			.77
8F	-1174.94	-2.24 (-1.32)		55.27 (5.02)*	77.85 (4.21)*	-8.45 (-0.60)	28.36 (1.34)	75.26 (4.29)*		.73
9-11F	-566.38	-1.06 (-0.63)		46.17 (4.21)*	65.89 (3.57)*	-13.97 (-0.99)	35.80 (1.69)*	72.38 (4.14)*		.68
12F	482.05	0.56 (0.33)		31.29 (2.85)*	52.51 (2.85)*	-19.23 (-1.34)	37.02 (1.75)*	68.49 (3.91)*		.62
13-15F	767.23	1.76 (1.02)		28.72 (2.57)*	56.54 (3.01)*	-21.52 (-1.49)	41.08 (1.91)*	63.81 (3.58)*		.61
16F	305.20	3.51 (1.83)*		34.91 (2.81)*	87.66 (4.20)*	-1.96 (-0.12)	35.64 (1.49)	62.41 (3.15)*		.67

t-values are shown below coefficients (critical value = 1.69; 38 d.f.)

* significant at the 5% or at a higher level with the correct sign

● significant at the 5% or at a higher level with the wrong sign

TABLE 4.2

MODEL II REGRESSION EQUATIONS FOR MALES

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	D	$\bar{R}^2$
0-7M	8411.54	0.18 (0.06)	-114.82 (-3.49)*	8.92 (0.29)	27.58 (0.70)	-79.93 (-2.02)*	-63.32 (-1.19)		1492.10 (0.77)	.77
8M	-404.03	2.03 (0.71)		57.19 (0.79)	110.64 (3.77)*	22.21 (0.96)	40.04 (1.15)	91.61 (2.99)*	3944.77 (0.60)	.76
9-11M	-2402.97	2.09 (0.79)		91.69 (5.51)*	179.66 (2.30)*	13.61 (0.63)	41.20 (1.28)	107.94 (4.02)*	-9091.18 (-1.05)	.75
12M	-1181.60	1.93 (0.75)		79.18 (4.76)*	94.41 (3.33)*	63.19 (1.35)	48.39 (1.51)	111.24 (4.20)*	-5511.90 (-1.32)	.72
13-15M	1822.61	0.48 (0.16)		63.86 (3.43)*	83.45 (2.64)*	-7.10 (-0.29)	-9.41) (-0.12)	146.08 (4.07)*	5158.90 (0.62)	.66
16M	3290.70	-0.53 (-0.20)		53.56 (3.10)*	103.73 (3.44)*	-20.40 (-0.94)	16.13 (0.49)	93.08 (2.49)*	12335.26 (2.35)*	.71

t-values are shown below coefficients (37 d.f.)

* significant at the 5% or at a higher level with the correct sign

● significant at the 5% or at a higher level with the wrong sign

TABLE 4.2a

MODEL II REGRESSION EQUATIONS FOR FEMALES

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	D	$\bar{R}^2$
0-7F	5573.02	-2.76 (-1.76)●	-109.49 (-5.31)*	-9.78 (-0.60)	10.70 (0.50)	-63.72 (-2.92)*	-44.30 (-1.55)		5998.84 (2.81)*	.81
8F	-326.19	-1.94 (-1.19)		-17.90 (-0.50)	74.35 (4.19)*	2.95 (0.20)	23.70 (1.16)	66.47 (3.85)	7026.70 (2.15)*	.76
9-11F	-1013.42	-0.64 (-0.38)		41.86 (3.72)*	89.44 (3.65)*	-16.10 (-1.15)	44.32 (2.04)*	65.89 (3.69)*	-3962.42 (-1.43)	.69
12F	-100.02	0.75 (0.45)		27.95 (2.50)*	53.81 (2.94)*	1.63 (0.08)	31.18 (1.45)	75.95 (4.16)*	-2154.05 (-1.31)	.63
13-15F	414.82	1.56 (0.92)		24.11 (2.12)*	51.41 (2.75)*	-13.84 (-0.92)	71.41 (2.49)*	61.08 (3.48)*	-4455.16 (-1.56)	.63
16F	147.30	3.65 (1.89)*		35.55 (2.84)*	86.04 (4.09)*	-5.13 (-0.31)	38.77 (1.59)	76.60 (2.89)*	-3553.30 (-0.81)	.66

t-values are shown below coefficients (37 d.f.)

* significant at the 5% or at a higher level with the correct sign

● significant at the 5% or at a higher level with the wrong sign

TABLE 4.3

MODEL III REGRESSION EQUATIONS FOR MALES

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	D	$\bar{R}^2$
0-7M	4681.03	-1.73 (-0.74)	-75.63 (-2.66)*	11.92 (0.47)	37.34 (1.21)	-56.17 (-1.78)*	-2.76 (-0.07)		1567.15 (1.05)	
		PF	PN	PU						
		7.96 (0.43)	8.04 (0.86)	17.53 (4.57)*						.88
8M	-466.74	-0.35 (-0.15)		43.44 (0.72)	97.77 (3.25)*	12.31 (0.53)	67.94 (2.40)*	70.42 (2.62)*	3380.85 (0.58)	
		PF	PN	PU						
		-8.10 (-0.43)	13.24 (1.34)	13.73 (3.43)*						.86
9-11M	-1453.89	-0.39 (-0.17)		71.03 (2.93)*	127.07 (1.94)*	4.51 (0.22)	66.60 (2.46)*	82.37 (3.16)*	-4583.55 (-0.66)	
		PF	PN	PU						
		-5.62 (-0.32)	11.80 (1.31)	13.44 (3.70)*						.85

t-values are shown below coefficients (34 d.f.)

* significant at the 5% or at a higher level with the correct sign

● significant at the 5% or at a higher level with the wrong sign

TABLE 4.3

MODEL III REGRESSION EQUATIONS FOR MALES
(Continued)

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	D	$\bar{R}^2$
12M	-681.28	-0.34 (-0.15)		64.61 (2.58)*	83.98 (2.83)*	37.55 (0.94)	74.27 (2.68)*	93.82 (3.54)*	-3621.64 (-1.02)	
		PF	PN	PU						
		-10.58 (-0.58)	13.74 (1.46)	11.89 (3.14)*						.82
13-15M	1569.46	-1.63 (-0.61)		54.95 (1.94)*	79.15 (2.33)*	-7.94 (-0.32)	73.07 (1.06)	125.02 (3.68)*	-732.88 (-0.10)	
		PF	PN	PU						
		-27.12 (-1.31)	17.53 (1.65)	10.92 (2.54)*						.78
16M	3772.29	-2.61 (-1.08)		41.47 (1.57)	91.34 (2.89)*	-23.83 (-1.01)	39.04 (1.32)	114.38 (2.83)*	7379.64 (1.48)	
		PF	PN	PU						
		-30.50 (-1.59)	13.78 (1.30)	8.97 (2.23)*						.79

t-values are shown below coefficients (34 d.f.)

* significant at the 5% or at a higher level of significance with the correct sign

● significant at the 5% or at a higher level of significance with the wrong sign

TABLE 4.3a

MODEL III REGRESSION EQUATIONS FOR FEMALES

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	D	$\bar{R}^2$
0-7F	1124.88	-1.06 (-0.86)	-67.50 (-3.74)*	-1.41 (-0.11)	17.50 (1.11)	-30.98 (-1.86)*	-24.00 (-1.14)		5980.98 (3.67)*	
		PF	PN	PU						
		52.65 (5.61)●	-11.67 (-2.41)*	7.31 (3.56)*						.91
8F	-604.92	-1.01 (-0.72)		-32.82 (-1.03)	39.38 (2.17)*	-12.03 (-0.87)	10.94 (0.63)	26.13 (1.60)	4969.20 (1.79)*	
		PF	PN	PU						
		43.99 (4.01)●	-9.14 (-1.62)	9.34 (3.88)*						.86
9-11F	-1000.51	-0.38 (-0.27)		10.39 (0.68)	47.99 (1.93)*	-25.78 (-2.00)●	34.69 (1.91)*	27.72 (1.71)*	-2345.43 (-1.05)	
		PF	PN	PU						
		37.10 (3.33)●	-4.06) (-0.69)	10.74 (4.55)*						.83

t-values are shown below coefficients (34 d.f.)

* significant at the 5% or at a higher level with the correct sign

● significant at the 5% or at a higher level with the wrong sign

TABLE 4.3a

MODEL III REGRESSION EQUATIONS FOR FEMALES
(Continued)

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	D	$\bar{R}^2$
12F	8.17	0.54 (0.38)		2.18 (0.14)	27.81 (1.47)	-18.18 (-0.95)	35.12 (1.98)*	37.73 (2.11)*	-1035.16 (-0.79)	
		PF	PN	PU						
		26.48 (2.33)●	0.10 (0.02)	10.60 (4.37)*						.79
13- 15F	858.88	1.95 (1.32)		-18.56 (-1.04)	10.91 (0.52)	-29.00 (-2.04)●	66.74 (2.82)*	17.40 (0.96)	-5157.46 (-1.75)●	
		PF	PN	PU						
		33.71 (2.61)●	-9.52 (-1.38)	10.75 (4.34)*						.78
16F	1319.44	3.12 (1.62)		-4.25 (-0.20)	50.92 (2.03)*	-23.89 (-1.33)	37.52 (1.59)	41.72 (1.51)	-3082.64 (-0.74)	
		PF	PN	PU						
		10.81 (0.69)	-8.01 (-1.03)	10.00 (3.12)*						.73

t-values are shown below coefficients (34 d.f.)

* significant at the 5% or at a higher level of significance with the correct sign

● significant at the 5% or at a higher level of significance with the wrong sign

a. The KL Variable. The KL variable has the correct sign and is significant only for the 16 labor type of females, and only in the first two models. In those same two models it has the wrong sign with a high t-value for the 0-7 female labor type. The influence of left-out variables is an acceptable explanation for these results in that the variable is insignificant for every equation in the third model.[4]

b. The Educational Categories. The 8, 9-11, and 16 educational categories of Model I are significant in all equations for all labor types except for 0-7 of both males and females, and in each instance they carry the correct sign. In these instances, 0-7 is the omitted category, and therefore a positive sign is predicted. All three categories in Model III, which provides the greatest explanation for area earnings as judged by the $\overline{R}^2$ for all equations, are significant

---

[3]It is to be noted that the 0-7 category enters only the equations for the 0-7 labor type. As is evident from the Tables, it is the omitted category in the equations for all other labor types. As is also evident from the Tables, the 16 category is omitted in the equations for the 0-7 labor type.

[4]Rao and Miller indicate that "When an independent variable in the true relation is omitted, the regression coefficients ... are biased. ... When a variable from the true relation is left out, a part of its influence in explaining the movements of the dependent variable is captured by the other independent variables. The relative share of each included variable in capturing the influence of the left-out variable is given by the auxiliary regression coefficients." Potluri Rao and LeRoy Miller, op. cit., p. 32. Thus, when an included independent variable has a strong relationship with one or more left-out variables in the true relation, both the size and significance of its regression coefficient will be affected. Since the regression coefficients of KL became insignificant when several variables were added to form Model III, it is possible that the significant coefficients of KL in Model I were capturing part of the influence of one or more of the significant additional variables of Model III.

in fewer of the female equations. By contrast, only the 8 category decreases in significance for the male equations, and then only for the 8 and 16 labor types.

Those categories which are significant in the equations for the 0-7 labor type of both males and females also carry the correct sign. In these instances, 16 is the omitted category, and therefore a negative sign is predicted.

In fact, all educational categories in all equations for all three models carry the correct sign, except for the 12 category in the equations for the 9-11 and 13-15 female labor types of Model III. There is no clear economic reason for these two categories to carry an incorrect sign. It is possible that it indicates a swamping of the complement relation by the substitute relation.

c. <u>The Demand Variable</u>. The general insignificance of the demand variable D, is indicated in the identical or near-identical $\bar{R}^2$ for all corresponding equations of Models I and II.[5] It is significant only for the 16 male labor type in Model II, for the 0-7 and 8 female types in both Model II

---

[5]Its lack of significance for many equations in Model II could have been a basis for dropping the demand variable for those equations in Model III, but it was retained in all equations of Model III on the basis of the argument expressed by Rao and Miller: "When the theory unambiguously states that a variable is a specified explanatory variable then of course, it should not be omitted even though it might appear superfluous." Potluri Rao and LeRoy Miller, op. cit., p. 37. In fact, this decision had no practical consequence. Regression equations were calculated both with and without the demand variable for Model III, and neither the coefficients nor the signs of the variables were substantially different in the two sets of equations. The $\bar{R}^2$ differed significantly only for those equations where D is significant.

and Model III, and for the 13-15 female labor type in Model III. Only in this last instance does it carry an incorrect sign.

The insignificance of this variable in every male equation of Model III and the general significance of the educational categories in those same equations supports the use of a marginal product demand function for the male labor types. A similar conclusion does not follow for the female labor types in that the demand variable is significant with the correct sign for both the 0-7 and 8 female types and the educational categories are generally insignificant in the female equations. It is possible that a more complex demand function, which would identify the complement and substitute relations, could explain the contradictory signs of the several female types, but the general insignificance of the cooperating labor types suggests that males and females require separate theories to explain the determinants of their earnings.[6]

d. The Percent Nonwhite. The PN variable is inconsistent in sign from equation to equation, and in the one instance

---

[6]A more complex demand function for a labor type as identified by educational category was developed in George E. Johnson, "The Demand for Labor by Educational Category," Southern Economic Journal, 37, October 1970, pp. 190-204. Johnson developed the function on the premise that the price of labor quality influences the hiring decisions of individual business firms. According to his theory the slope of the aggregate demand schedule depends on the degree to which employers can substitute less for more educated workers and on differences in substitute possibilities across firms. Basing his analysis on states of the union, he found the elasticity of substitution between college and other labor to be 1.3.

where it is significant, for the 0-7 female labor type, it carries a negative sign. According to the theory, the sign was predicted upon the relationship of nonwhite to white productivity. Since the variable was significant in only this one instance, no general conclusions can be derived.

The negative sign affirms that the earnings of females in the lowest educational category are inversely related to the percentage of nonwhites in the area, and this is the most likely type to be affected in this manner if the number of nonwhites is a satisfactory index of the number of female nonwhites in an area. Both because of limited education and occupational discrimination, nonwhite women are forced to seek jobs at the lowest skill level. These jobs are both least organized and most open to wage discrimination. Furthermore, in contrast with white women, the nonwhite women must work in order to meet minimum family needs. This necessity further reduces an already small bargaining power.[7]

e. The Percent Female. The next variable to be considered, PF, is inconsistent in sign for the male equations but is not significant in any of them. By contrast, it is highly significant in every female equation, except for labor

---

[7] In his analysis of Census data, Alan Batchelder noticed that American Negro women have always borne exceptionally heavy family responsibility. "In 1910 there were ... 67 employed Negro women for every 100 Negro men. Even in 1959, only 8 percent of white families but 21 percent of nonwhite families were headed by women. Three-fourths of these nonwhite families were poor in 1959." Alan Batchelder, "Poverty: The Special Case of the Negro," Perspectives On Poverty and Income Distribution, James G. Scoville, ed., D.C. Heath & Co., Mass., 1971, p. 120.

type 16, and in each case it carries a positive sign. The expected sign was negative, primarily because of the assumption that women are weaker resources than men. This was contradicted by the positive sign. Another important assumption, that all categories vary proportionately with a change in PF, was not proved to be totally unacceptable for empirical purposes, but the uneven impact of PF on the two sexes; i.e., significant for females but insignificant for men, places into question the use of sex as a labor-quality index.

The alternate theoretical approach, as expressed in Eq. 2.11, assumed no specific relationship between male and female productivity, and a change in PF required no specific limitations on the male-female ratios in the separate categories. The sign was therefore contingent upon the actual productivity of females relative to males and the male educational category allowed to vary with PF. A positive sign, interpreted according to this theory, indicates that the impact of the percent female on the earnings of all but the highest female labor type is greater than either of the omitted categories. But the relative size of the coefficients indicates that the relationship between PF and the labor types is not consistent.

Within this same theoretical framework, there is another way that the coefficient of PF might be interpreted. Equation 2.10 defines the educational categories in terms of the male labor force with the percent female being treated as a cooperating factor. If that equation is rewritten to define the educational categories in terms of the female labor force, a new understanding emerges. This becomes most clear when a category is omitted and an equation comparable to Eq. 2.11 is derived:

$$Q = \beta_1(K/L) + \beta_2 PF + (\beta_3-\beta_2)L_2/L + \ldots + (\beta_7-\beta_2)L_6/L + (\beta_8-\beta_2)L_m/L. \qquad \text{Eq. 4.1}$$

In this equation the percent male is treated as a cooperating factor and the coefficient of the percent female is the coefficient of the omitted category. Therefore the coefficient of PF is expected to be positive.

This interpretation seems acceptable because of the very high correlation between the corresponding male and female categories. However, the size of the coefficient for PF in the equations for the different labor types encourages hesitation in the acceptance of this interpretation. If the coefficient of PF is actually the coefficient of the omitted category, why is there no consistent pattern for the types?

But there is still another way that the coefficient of PF might be nterpreted. In the light of the relatively insignificant and inconsistent performance of the demand variable, D, it is possible that it was not an index of demand at all, that it captured a relationship which was not articulated in the theory and which is not clear from the results. If this is true, the original interpretation of PF as an index of labor force quality may simply be incorrect. It may have served in the regressions as an index of demand for females. It may have captured the response of the female labor types to the wages offered. The observations of Herman P. Miller give support to this as a possibility:

> ... the proportion of wives in the paid labor force increased from 19 percent in 1950 to 26 percent in 1960. Many factors contributed to this change. One major factor was the relatively full employment situation throughout the postwar period with the concomitant high demand for labor. In addition, the increased social acceptance of women as workers encouraged women to seek jobs, and the widespread use of labor-saving devices made it possible for many to take on the added burdens of a paid job in addition to their other duties. Indeed the general attitude appears to be that housekeeping is not the full-time job it was once thought to be, especially when the children are of school age. Other factors underlying the general emergence of women as paid workers are the changing needs of the economy and the relatively high increase in the educational attainment of women, which enabled many to qualify for employment.[8]

In summary, the coefficient of PF can be explained in at least three possible ways: as the net result of the productivity of females in the aggregate to the productivity of the omitted category; as the effect of the omitted category in the female equations; or as an index of area demand for females.

f. The Percent Under Union Contract. As stated previously, PU has both the correct sign and possesses a high degree of significance in every equation of every model. This indicates that the earnings of all labor types, including college graduates, are related to area unionization, and it supports previous findings on the importance of unions as a force in wage determination.

---

[8]Herman P. Miller, Income Distribution in the United States, (1960 Census Monograph), U.S. Government Printing Office, 1966, pp. 50 ff.

2. Relative Size of Coefficients and Other Relationships

Several patterns become evident in a comparison of the relative size of coefficients in all equations simultaneously. Three variables dominate Model I which is listed in Table 4.1, the educational categories, 8, 9-11, and 16. The coefficients for the 8 and 9-11 categories are largest for labor types identified by the lowest level of education, 0-7. They are successively smaller for each labor type identified by a higher level of education.[9] Thus, the percentage of the labor force having completed eighth grade is related most strongly to the earnings of the 9-11 labor type and least to labor type 16 in the male equations. A similar statement can be made for the 9-11 labor type, which indicates a greater complementarity between similar labor types. The coefficients of the 16 category support this conclusion, since they are largest for the labor types identified by high educational categories and least for those types identified by the low educational categories.

The same reasoning does not hold for the female equations listed in Table 4.1a. The coefficients for the 16 category are largest for the lowest labor type and are successively

---

[9]The relative effect on different categories is even more pronounced if the coefficients are considered from the viewpoint of a percentage change in average earnings. Since each labor type is identified with successively higher average earnings, the same size coefficient for every labor type would indicate a successivley smaller effect on each. For example, if the coefficient were 100 for every labor type, it would indicate a 10% change in average earnings of $1000 but only 1.25% change in average earnings of $8000. Since, in fact, the coefficient decreases for each higher labor type, the impact is even smaller than suggested by the coefficients.

smaller through the highest. The coefficients for the 8 and 9-11 categories are large, become smaller, and then become large again.

A comparison with Model III suggests that little, if any, conclusions should be drawn from the above. Fewer co-efficients for the female equations are significant, and these follow no pattern. The coefficients for the male labor types also change in significance from model to model, except for the 16 category.

As stated previously, the PF variable is insignificant for all male equations and significant with a positive sign for all female equations. The value of the coefficient follows a pattern similar to other variables in that it is largest for the 0-7 labor type and successively smaller for the higher labor types with the exception of labor type 16. This can be interpreted in terms of substitute and complement relations with reference to the theory as developed. But if PF is an index of demand for females as was suggested, this pattern has no special meaning. There is no inherent reason to expect the relationship to be of any particular magnitude for the separate labor types, so the pattern or deviation from a pattern would have no special interpretation.

The coefficients of PU indicate an uneven effect on different labor types and a different pattern for males and females. The male coefficient is largest for the 0-7 labor type and it decreases in size for every successively higher type. By contrast, the female coefficient is smallest for the 0-7 type, and it has no increasing or decreasing pattern. The values of the four highest female categories do not vary greatly, but they decrease in effect from the lowest to the highest when calculated as a percentage change in average earnings. According to this mode of measuring change, the three lowest female labor types are affected more by area unionization than any of the male types. This measurement also indicates that area unionization generally has a greater impact on females than it does on males, and it has its greatest impact on the lower labor types.

a. Patterns in $\overline{R}^2$. The general pattern of $\overline{R}^2$ in all three models indicates that the variables in each model explain most of the variance in earnings for the lowest labor type, 0-7, and successively less for each higher type, except for labor type 16. This result is consistent with an expectation that lower labor types are relatively immobile both because of limited job opportunities in other areas and because of a lack of knowledge as to the opportunities available. The labor types identified by the higher educational categories should be most independent of area characteristics for the opposite reason. It is more likely for a college graduate to be aware of job opportunities in other areas and to possess attitudes conducive to mobility. As a result, their earnings would have a greater tendency to be related to national forces than other labor types.

The relative size of $\overline{R}^2$ for corresponding equations indicates that Model III explains substantially more of the earnings variance than either of the other two models.

3. Regional Differentials

Tables 4.4, 4.5, and 4.6 list the results of adding a dummy variable (North = 0; South = 1) to every equation in all three models. Noticably, the variable has the expected sign for every equation in every model. It is also significant for every female equation in every model, except for the 16 labor type of Model III. In the male equations, it is significant only for the 0-7 and 12 labor types of Model II.

With the addition of the dummy variable, the coefficients in every equation for every model generally decrease in value. In every model, some variables that had been significant become insignificant and several variables that were insignificant become significant. For example, the demand variable, D, had been significant with the correct sign for both 0-7 and 8 female labor types. In Model II of the new equations it is no longer significant for either of these labor types and it has become significant with the wrong sign for both the 12 male and 13-15 female labor types. But in the Model III it is again significant for the 0-7 female labor type, is no longer significant for the 12 male type, and remains significant with the wrong sign for the 13-15 female type.

As indicated by the $\overline{R}^2$ for all equations relative to corresponding equations in the other two models, Model III provides the best explanation for earnings differentials. An analysis of this model reveals that the major effect of

TABLE 4.4

MODEL I REGRESSION EQUATIONS FOR MALES WITH REGIONAL VARIABLE

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	RD	$\bar{R}^2$
0-7M	8658.77	0.01 (0.00)	-87.83 (-2.97)*	-7.97 (-0.26)	11.02 (0.28)	-80.70 (-2.10)*	-64.30 (-1.25)		-330.33 (-1.66)	.78
8M	597.60	1.66 (0.64)		72.74 (3.13)*	83.70 (2.52)*	1.23 (0.05)	36.67 (1.12)	85.45 (3.05)*	-310.79 (-1.66)	.77
9-11M	823.80	1.55 (0.61)		69.82 (3.07)*	78.56 (2.41)*	0.35 (0.01)	34.65 (1.08)	91.64 (3.33)*	-260.84 (-1.42)	.76
12M	1676.11	1.88 (0.73)		58.47 (2.54)*	66.68 (2.02)*	-3.96 (-0.17)	39.65 (1.22)	99.12 (3.56)*	-221.61 (-1.19)	.72
13-15M	2522.74	1.30 (0.46)		46.48 (1.83)*	62.43 (1.72)*	-12.97 (-0.49)	26.87 (0.75)	125.05 (4.08)*	-183.86 (-0.90)	.66
16M	3855.26	1.13 (0.41)		34.24 (1.40)	67.54 (1.93)*	-16.17 (-0.64)	-6.44 (-0.19)	150.75 (5.09)*	-93.44 (-0.47)	.66

t-values are shown below coefficients (37 d.f.)

* significant at the 5% or at a higher level with the correct sign

● significant at the 5% or at a higher level with the wrong sign

TABLE 4.4a

MODEL I REGRESSION EQUATIONS FOR FEMALES WITH REGIONAL VARIABLE

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	RD	$\bar{R}^2$
0-7F	5974.51	-3.52 (-2.43)●	-54.54 (-3.51)*	-25.73 (-1.61)	-4.05 (-0.19)	-81.92 (-4.06)*	-48.26 (-1.79)*		-374.05 (-3.58)*	.83
8F	807.34	-1.99 (-1.31)		24.70 (1.82)*	44.30 (2.27)*	-28.23 (-2.01)●	17.41 (0.90)	59.09 (3.59)*	-356.58 (-3.25)*	.79
9-11F	1372.70	-0.82 (-0.53)		16.27 (1.19)	33.06 (1.69)*	-33.32 (-2.36)●	25.09 (1.30)	56.56 (3.43)*	-348.81 (-3.17)*	.75
12F	2036.81	0.75 (0.47)		7.31 (0.51)	26.19 (1.28)	-34.75 (-2.35)●	28.43 (1.41)	55.81 (3.23)*	-279.68 (-2.43)*	.67
13-15F	2480.22	1.98 (1.23)		2.30 (0.16)	27.54 (1.34)	-38.61 (-2.60)●	31.62 (1.56)	49.84 (2.87)*	-308.14 (-2.66)*	.68
16F	1999.62	3.72 (2.04)*		8.78 (0.54)	58.98 (2.53)*	-18.87 (-1.12)	26.28 (1.14)	48.59 (2.47)*	-304.80 (-2.32)*	.71

t-values are shown below coefficients (37 d.f.)

* significant at the 5% or at a higher level with the correct sign

● significant at the 5% or at a higher level with the wrong sign

TABLE 4.5

MODEL II REGRESSION EQUATIONS FOR MALES WITH REGIONAL VARIABLE

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	D	$\bar{R}^2$
0-7M	8788.52	0.52 (0.21)	-101.76 (-3.10)*	-4.35 (-0.14)	10.53 (0.27)	-84.03 (-2.18)*	-56.60 (-1.09)		1865.34	
		RD								
		-352.12 (-1.76)*								.78
8M	1235.54	2.19 (0.79)		34.80 (0.48)	81.95 (2.43)*	4.83 (0.19)	31.18 (0.90)	78.42 (2.53)*	3576.75 (0.55)	
		RD								
		-307.12 (-1.62)								.77
9-11M	-785.56	2.15 (0.81)		71.23 (3.12)*	144.81 (1.76)*	0.52 (0.02)	34.08 (1.06)	96.36 (3.43)*	-7629.67 (-0.88)	
		RD								
		-239.57 (-1.29)								.76

t-values are shown below coefficients (36 d.f.)

* significant at the 5% or at a higher level with the correct sign

● significant at the 5% or at a higher level with the wrong sign

TABLE 4.5

MODEL II REGRESSION EQUATIONS FOR MALES WITH REGIONAL VARIABLE
(Continued)

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	D	$\bar{R}^2$
12M	-35.49	2.25		51.54	66.20	68.23	39.06	97.12	-7861.45	
		(0.90)		(2.27)*	(2.07)*	(1.49)	(1.24)	(3.59)*	(-1.83)●	
		RD								
		-330.83								.74
		(-1.47)*								
13-15M	2876.39	0.58		47.86	65.97	-17.65	-16.79	138.08	5358.22	
		(0.19)		(1.86)*	(1.78)*	(-0.64)	(-0.22)	(3.73)*	(0.64)	
		RD								
		-187.31								.66
		(-0.91)								
16M	4250.27	-0.52		39.49	89.32	-30.65	12.17	80.98	13180.92	
		(-0.20)		(1.71)*	(2.62)*	(-1.25)	(0.36)	(2.04)*	(2.47)*	
		RD								
		-173.17								.71
		(-0.92)								

t-values are shown below coefficients (36 d.f.)

* significant at the 5% or at a higher level with the correct sign

● significant at the 5% or at a higher level with the wrong sign

TABLE 4.5a

MODEL II REGRESSION EQUATIONS FOR FEMALES WITH REGIONAL VARIABLE

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	D	$\bar{R}^2$
0-7F	5879.66	-3.05 (-2.07)●	-77.73 (-3.32)*	-21.98 (-1.37)	-1.85 (-0.09)	-73.66 (-3.52)*	-44.58 (-1.66)		3081.51 (1.32)	
		RD								
		-292.26 (-2.42)*								.84
8F	1183.09	-1.81 (-1.21)		-23.96 (-0.73)	45.69 (2.40)*	-17.80 (-1.17)	15.36 (0.81)	54.69 (3.35)*	5012.49 (1.63)	
		RD								
		-315.26 (-2.86)*								.81
9-11F	1133.92	-0.70 (-0.45)		16.61 (1.20)	41.87 (1.47)	-32.90 (-2.30)●	28.21 (1.36)	55.48 (3.29)*	-1182.51 (-0.43)	
		RD								
		-329.85 (-2.76)*								.74

t-values are shown below coefficients (36 d.f.)

* significant at the 5% or at a higher level with the correct sign

● significant at the 5% or at a higher level with the wrong sign

TABLE 4.5a

MODEL II REGRESSION EQUATIONS FOR FEMALES WITH REGIONAL VARIABLE
(Continued)

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	D	$\bar{R}^2$
12F	1497.53	0.89		6.30	28.84	-17.95	24.55	62.26	-1632.54	
		(0.55)		(0.44)	(1.40)	(-0.82)	(1.19)	(3.39)*	(-1.03)	
		RD								
		-262.03								.67
		(-2.25)*								
13-15F	2133.11	1.77		-2.56	22.13	-30.92	62.40	46.97	-4530.38	
		(1.13)		(-0.18)	(1.09)	(-2.04)●	(2.34)*	(2.77)*	(-1.72)●	
		RD								
		-310.16								.70
		(-2.75)*								
16F	2404.37	3.64		3.95	55.51	-19.16	22.26	35.11	2803.55	
		(1.97)*		(0.21)	(2.28)*	(-1.13)	(0.91)	(1.12)	(0.55)	
		RD								
		-355.20								.70
		(-2.21)*								

t-values are shown below coefficients (36 d.f.)

* significant at the 5% or at a higher level with the correct sign

● significant at the 5% or at a higher level with the wrong sign

TABLE 4.6

MODEL III REGRESSION EQUATIONS FOR MALES WITH REGIONAL VARIABLE

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	D	$\bar{R}^2$
0-7M	5112.27	-1.46 (-0.62)	-74.97 (-2.65)*	8.01 (0.31)	30.35 (0.97)	-59.61 (-1.89)*	-4.40 (-0.11)		1752.24 (1.17)	
		RD	PF	PN	PU					
		-193.92 (-1.16)	7.20 (0.39)	11.46 (1.17)	15.78 (3.85)*					.88
8M	368.23	-0.02 (-0.01)		28.03 (0.47)	88.76 (2.89)*	8.03 (0.35)	62.36 (2.20)*	67.77 (2.54)*	4304.92 (0.75)	
		RD	PF	PN	PU					
		-211.27 (1.28)	-10.42 (-0.56)	17.60 (1.70)•	11.58 (2.69)*					.87
9-11M	-845.97	-0.31 (-0.14)		66.45 (2.67)*	114.42 (1.70)*	0.76 (0.04)	63.67 (2.33)*	80.67 (3.07)*	-3855.13 (-0.55)	
		RD	PF	PN	PU					
		-140.02 (-0.87)	-6.58 (-0.37)	14.36 (1.51)	12.19 (3.06)*					.85

t-values are shown below coefficients (33 d.f.)

* significant at the 5% or at a higher level with the correct sign

• significant at the 5% or at a higher level with the wrong sign

TABLE 4.6

MODEL III REGRESSION EQUATIONS FOR MALES WITH REGIONAL VARIABLE
(Continued)

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	D	$\bar{R}^2$
12M	-354.77	0.00 (0.00)		56.80 (2.22)*	74.51 (2.45)*	44.90 (1.12)	69.02 (2.48)*	91.15 (3.45)*	-4991.34 (-1.36)	
		RD	PF	PN	PU					
		-212.09 (-1.24)	-10.52 (-0.58)	16.84 (1.75)●	10.06 (2.49)*					.83
13-15M	1931.85	-1.52 (-0.56)		51.43 (1.75)*	74.17 (2.10)*	-10.94 (-0.43)	70.22 (1.01)	124.14 (3.61)*	-680.72 (-0.09)	
		RD	PF	PN	PU					
		-108.02 (-0.57)	-27.82 (-1.32)	19.46 (1.73)●	9.93 (2.12)*					.77
16M	4017.75	-2.56 (-1.04)		39.09 (1.42)	88.89 (2.69)*	-26.05 (-1.06)	37.65 (1.25)	112.94 (2.73)*	7593.70 (1.50)	
		RD	PF	PN	PU					
		-69.70 (-0.39)	-30.86 (-1.59)	14.86 (1.34)	8.35 (1.91)*					.79

t-values are shown below coefficients (33 d.f.)

*significant at the 5% or at a higher level with the correct sign

● significantat the 5% or at a higher level with the wrong sign

TABLE 4.6a

MODEL III REGRESSION EQUATIONS FOR FEMALES WITH REGIONAL VARIABLE

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	D	$\bar{R}^2$
0-7F	1556.49	-1.25 (-1.04)	-54.97 (-2.93)*	-6.99 (-0.54)	11.18 (0.71)	-37.99 (-2.29)*	-26.70 (-1.30)		4478.13 (2.51)*	
		RD	PF	PN	PU					
		-168.96 (-1.81)*	50.19 (5.46)●	-9.36 (-1.93)*	6.32 (3.07)*					.92
8F	41.56	-0.87 (-0.67)		-28.33 (-0.96)	28.71 (1.65)	-20.72 (-1.55)	7.10 (0.44)	25.00 (1.64)	3749.41 (1.42)	
		RD	PF	PN	PU					
		-235.51 (-2.47)*	42.10 (4.10)●	-4.66 (-0.84)	7.48 (3.16)*					.88
9-11F	82.26	-0.25 (-0.19)		1.57 (0.11)	24.45 (0.97)	-32.78 (-2.64)●	24.57 (1.40)	26.12 (1.72)*	-699.95 (-0.32)	
		RD	PF	PN	PU					
		-238.83 (-2.41)*	36.90 (3.54)●	-1.07 (-0.19)	8.87 (3.79)*					.85

t-values are shown below coefficients (33 d.f.)

* significant at the 5% or at a higher level with the correct sign

● significant at the 5% or at a higher level with the wrong sign

TABLE 4.6a

MODEL III REGRESSION EQUATIONS FOR FEMALES WITH REGIONAL VARIABLE
(Continued)

TYPE	INTERCEPT	KL	0-7	8	9-11	12	13-15	16	D	$\bar{R}^2$
12F	661.30	0.71 (0.51)		-3.64 (-0.23)	19.15 (1.01)	-25.11 (-1.33)	31.61 (1.82)*	35.12 (2.02)*	-836.05 (-0.65)	
		RD	PF	PN	PU					
		-178.24 (-1.76)*	25.48 (2.30)●	3.23 (0.54)	9.02 (3.59)*					.81
13-15F	1522.67	2.14 (1.49)		-24.35 (-1.39)	2.50 (0.12)	-34.67 (-2.47)●	61.03 (2.65)*	16.06 (0.92)	-4899.83 (-1.72)●	
		RD	PF	PN	PU					
		-196.56 (-1.89)*	31.92 (2.55)●	-5.69 (-0.82)	8.96 (3.48)*					.80
16F	2358.02	3.03 (1.58)		-13.03 (-0.58)	42.21 (1.62)	-27.53 (-1.52)	30.57 (1.27)	25.24 (0.82)	704.27 (0.13)	
		RD	PF	PN	PU					
		-207.59 (-1.19)	5.95 (0.37)	-4.68 (-0.57)	8.22 (2.34)*					.73

t-values are shown below coefficients (33 d.f.)

* significant at the 5% or at a higher level with the correct sign

● significant at the 5% or at a higher level with the wrong sign

the dummy variable in the male equations is to render the PN variable significant for three labor types. Since the $\overline{R}^2$ is relatively unchanged for corresponding equations in the model without the dummy variable and since the dummy variable is insignificant in all three equations, the dummy variable can be considered superfluous.

The major effect of the dummy variable for the female equations is to render the 9-11 category insignificant in the three equations wherein it had been significant. Since the dummy variable is significant and the $\overline{R}^2$ is increased for all but labor type 16, the previous significance of the 9-11 category can be considered the result of specification bias.

An overview of all models, both with and without the dummy variable, indicates that Model III without the dummy variable is generally superior for the male labor types, and that Model III with the dummy variable is generally superior for the female labor types.

4. A Modification of Model III

The regression equations of Model III indicate that some cooperating factors are important in determining the earnings of the male labor types in general. These results become the basis for a question which had not been asked previously: Is the quality of the cooperating factors, considered as a group, sufficient to explain the difference in earnings between areas for a labor type? Or is the composition of the cooperating labor force -- quantity and quality as indicated by the educational categories -- more important? A comparison can be made with a small modification of Model III:

$$e^{m}_{ij} = f(KL_j, L_{ij}, L^{i}_{j}, D^{m}_{ij}, PF_j, PN_j, PU_j) \quad \text{Eq. 4.2}$$

$$e^{f}_{ij} = f(KL_j, L_{ij}, L^{i}_{j}, D^{f}_{ij}, PF_j, PN_j, PU_j) \quad \text{Eq. 4.3}$$

where L^i = the average educational level of the male labor force after withdrawing the ith labor type, and all other variables remain as previously defined. The a priori expectation is that L^i is positive in sign. Since the quality of all other categories is accounted for with L^i, the sign of L_{ij} will depend on its quantity relative to the quantity of these cooperating labor types. Therefore, a negative sign is expected.

Table 4.7 lists the regressions which correspond to this empirical model. It is to be noted the $\bar{L}$ for all equations is the mean educational level of the total male labor force.[10] Six separate means were calculated as discussed above; i.e., for the remainder of the labor force after the proportion of the labor force corresponding to the labor type was withdrawn, but there was no need to distinguish these means separately in that all were very highly correlated with the mean educational level of the total male labor force. The correlation coefficient varied between .950 and .997.

---

[10] The correlation coefficient between the mean education of males and the mean education of females is .930, so the regression results would be almost identical if the female mean replaced the male mean in Eq. 4.3.

TABLE 4.7

MODIFIED MODEL III REGRESSION EQUATIONS FOR MALES

TYPE	INTERCEPT	KL	$\overline{L}$	L	PF	PN	PU	D	$\overline{R}^2$
0-7M	5963.82	-1.66 (-0.67)	-247.26 (-1.01)	-63.57 (-1.91)*	22.05 (1.12)	1.03 (0.11)	20.85 (5.82)*	1513.48 (0.95)	.86
8M	230.11	0.21 (0.09)	332.27 (2.62)*	-11.13 (-0.20)	-0.04 (-0.00)	4.20 (0.46)	17.20 (4.60)*	4232.57 (0.77)	.83
9-11M	-995.77	-1.17 (-0.49)	313.04 (3.06)*	62.17 (0.93)	3.23 (0.18)	0.32 (0.05)	18.40 (6.41)*	-2939.45 (-0.40)	.82
12M	-1143.43	-0.50 (-0.22)	615.23 (4.42)*	-26.88 (-0.63)	-12.17 (-0.64)	1.62 (0.26)	16.32 (6.03)*	-3254.71 (0.91)	.82
13-15M	-1670.71	-1.74 (-0.68)	978.61 (3.80)*	-38.67 (-0.50)	-29.21 (-1.43)	12.70 (1.77)●	12.68 (3.68)*	730.62 (0.10)	.79
16M	2923.83	-2.54 (-1.04)	608.28 (1.96)*	15.69 (0.28)	-32.39 (-1.62)	6.47 (0.79)	12.64 (4.26)*	6623.31 (1.32)	.78

t-values are shown below coeffients (37 d.f.)

* significant at the 5% or at a higher level with the correct sign

● significant at the 5% or at a higher level with the wrong sign

TABLE 4.7a

MODIFIED MODEL III REGRESSION EQUATIONS FOR FEMALES

TYPE	INTERCEPT	KL	$\overline{L}$	L	PF	PN	PU	D	$\overline{R}^2$
0-7F	2216.78	-0.71 (-0.55)	-204.97 (-1.64)	-72.51 (-3.34)*	61.71 (6.29)●	-14.48 (-2.99)*	8.82 (4.45)*	7387.06 (4.38)*	.89
8F	700.18	-0.76 (-0.52)	-18.54 (-0.24)	-71.45 (-2.49)*	51.39 (4.57)●	-14.50 (-2.71)*	11.16 (4.92)*	7023.62 (2.61)*	.84
9-11F	-1897.91	-0.30 (-0.20)	93.80 (1.36)	40.17 (1.78)●	46.88 (3.99)●	-1.85 (-0.41)	10.30 (5.41)*	-2181.08 (-0.93)	.80
12F	-1444.52	1.02 (0.72)	325.95 (3.62)*	-42.83 (-2.35)*	28.24 (2.41)●	2.25 (0.57)	9.71 (5.59)*	-597.48 (-0.48)	.78
13-15F	45.56	2.54 (1.59)	-31.66 (-0.26)	82.87 (2.71)●	46.47 (3.62)●	-4.03 (-0.84)	8.83 (3.80)*	-6799.57 (-2.50)●	.73
16F	2036.82	4.34 (2.10)*	-39.56 (-0.27)	50.04 (1.47)	20.85 (1.23)	-8.43 (-1.35)	9.46 (3.79)*	-3370.47 (-0.75)	.66

t-values are shown below coefficients (37 d.f.)

* significant at the 5% or at a higher level with the correct sign

● significant at the 5% or at a higher level with the wrong sign

Since they are all identical for empirical purposes, the mean for the total male labor force was used in all equations in order to simplify the computer program.

The results in general are consistent with the previous regressions. The PU variable is again significant with the correct sign in every equation for both males and females, the PF variable is not significant in any male equation and significant with a positive sign in every female equation except for labor type 16, the KL variable is insignificant in every equation except for labor type 16 of females, the educational category corresponding to the labor type, L, is significant with a negative sign for the 0-7 labor type of both males and females, and both D and PN are significant for the 0-7 and 8 female labor types.

There are few differences. The KL variable is significant for labor type 16 of females, which was stated above, the L variable is significant with inconsistent signs for all but the 16 female labor type, and PN is significant for 13-15 males. Almost all of these results can be found in one or more of the regressions for the models based on the composition of the cooperating labor force. By themselves they do not reveal any new information.

Of most importance, the mean education, $\overline{L}$, is significant for every male equation except labor type 0-7 and insignificant for every female equation except labor type 12. This result too is consistent with the previous model in that the cooperating factors were generally significant in the male equations and generally insignificant in the female equations. One new conclusion is suggested by the results here: that co-operating factors, whether measured by educational level or in terms of quantity and quality, are more important in the determination of earnings for the male labor types than the proportion of the labor force identified with the labor type itself. The opposite conclusion is suggested for females.

The difference in $\bar{R}^2$ for these equations relative to those containing the educational categories as specified in Model III is negligible, except for the two highest female labor types. Thus, the educational level of an area labor force can explain almost as much variance in earnings as the educational composition.

The equations were recalculated with a North-South dummy variable and the regressions remained relatively unchanged. The dummy variable was significant for the 8, 9, and 12 female labor types but the $\bar{R}^2$ remained relatively unchanged in all three equations. Again it carried the correct sign in all equations for both males and females, and the t-value, although not significant, was relatively large for both 0-7 and 13-15 females. These results are consistent with the previous models.

5. Conclusions and Policy Implications

This dissertation was designed with the intent of gaining insight into the economic forces which determine the earnings of a labor type, as defined by level of education, within a metropolitan area. The observed differential in such earnings between metropolitan areas provided the focus for the analysis.

Three models were developed. The first model was so proscribed that only a difference in factor proportions could account for an interarea earnings differential. The second model added a demand variable, and the third model added both the sex and race composition of the area labor force in addition to the degree of unionization. A North-South dummy variable was then added to account for possible regional variance. Finally, a modified model was developed in which the educational level of the area labor force replaced the labor force composition as a possible explanation of area earnings.

The theory was not fully supported in any one equation of any model, but the regression results for the male labor types were generally consistent with the proposition that cooperating factors within an area are significantly related to the earnings of a labor type. The regressions for the female labor types did not give evidence of such a relationship.

The general insignificance of the capital-labor ratio in every model tested was most surprising. There are several ways for this result to be interpreted. It is possible that the rewards from an increase in productivity due to a larger capital stock are completely absorbed by the owners of the capital. However, this is inconsistent with the results of other studies. It is also possible that these rewards are not allocated to either capital or labor in a consistent way, but this also is not consistent with the results of other studies. The third possibility seems most likely -- that the capital-labor ratio in manufacturing, at least as calculated here, is not a good proxy for the SMSA capital-labor ratio. It is highly probable that a better measure of this ratio would yield significant results.

The significance of the union variable in every equation of every model accents the importance of institutional forces in determining wages. Since the variable itself was an area percentage, the results are consistent with the long-held proposition that even non-union workers gain from union activity in an area.

In their analysis of the effects of unionization on the distribution of income, Johnson and Mieszkowski concluded that the gains of union labor are made at the expense of non-union workers, and not at the expense of earnings on capital.[11] If their conclusion is accepted in concert with the results of this study, then it is the non-union labor in the poorly organized areas that bear the ultimate burden of unionization in the nation.

A more important result of this study is the clear

---

[11]Harry G. Johnson and Peter Mieszkowski, "The Effects of Unionization on the Distribution of Income: A General Equilibrium Approach," Quarterly Journal of Economics, November 1970, 84, pp. 560-561.

indication that males and females are not related in the same way to area resources. Except for the labor types at both extremes of the queue, the male equations are generally consistent with the competitiveforces described by a neo-classical model. All models, including that which contained a simple index of cooperating factor quality, established a significant relationship between earnings and area resources for these male labor types. The female equations, by contrast, suggest a basic relationship with the forces of demand rather than a strong relationship with the supply of other factors. This result supports the claim of the women's liberation movement, that women are not treated on the same level as men in the marketplace.

A most disheartening conclusion was also expected: that the lowest labor type of both males and females is in the most defenseless position. The earnings of this type are related to the relative number of the type in the area labor force and to the competition from other labor types. The importance of unionization in determining the earnings of these people suggests that some type of institutional force should be considered to assist them. The results also indicate a negative relationship between the percentage of nonwhites in an area and the earnings of females of the lowest labor type. Discrimination against nonwhite women of this type can be inferred. Since this is the only type significantly related to the number of nonwhites in the area labor force, the results also suggest that wage discrimination is not practiced against the other nonwhite labor types. If discrimination is practiced against the other nonwhite labor types, it does not appear to be a systematic relationship in all areas. As pointed out earlier, the discrimination seems to be practiced by confining nonwhites to the lower skill jobs.

A regional differential was indicated for all but the highest female type and no differential at all for the male types. This is a very interesting result in that it is the only study in the large literature indicating a differential between the North and the South for women and not for men. The approach to earnings by labor type seems to have uncovered a phenomenon that has remained hidden in other studies.

a. Policy Implications

The main conclusions of this study can be summarized:

1. That the earnings of the male labor types in general are significantly related to the cooperating labor force in an area.

2. That area unionization provides a significant explanation for the interarea earnings differentials of all labor types, both male and female.

3. That the earnings of the lowest labor type are not significantly related to the quantity and quality of other area resources in general. They are related primarily to the extent of area unionization and to the percentage of the type relative to the area labor force. In addition,

earnings of the females are affected by the number of females relative to males in the labor force and nonwhite females suffer the burden of discrimination.

4. That a North-South differential exists for females but not for males.

The first conclusion is heartening in that it indicates competitive forces are at work and the differential due to different factor ratios can be removed through adequate mobility of factors. Policy decisions should encourage mobility by removing the obstructions due to discrimination, lack of knowledge, etc.

The second conclusion emphasizes the importance of institutional forces in resource allocation and factor earnings. To suggest a policy here in order to balance the influence of unionization, in the light of experience, is hazardous. Legislation for other reasons, primarily political, could and should be used to control union power. But legislation designed to offset union power and rechannel resources would be extremely difficult if not impossible to formulate, adopt, or administer. If a policy must be established, perhaps the most realistic is the encouragement of organization in the unorganized areas.

It was suggested by the regression results that unionization has the greatest effect on earnings of the lowest labor type. However, it should be noted that union success would depend upon the type of job and substitutability with other factors. Private household workers, for example, can be replaced in a variety of ways including a change of living habits to allow for more dirt in the house. In addition, this labor type may be very difficult to organize. A recent television special related a long history of union failures in attempts to organize day workers in Chicago.[12]

It must be recognized that laborers identified by a low level of education simply do not have much to sell in the labor market. Nor do they normally have the intelligence and skill to bargain shrewdly. Thus, it may be best simply to subsidize them in some way so that they may live a reasonably decent life.

The fourth conclusion indicates that more study is needed before any policy can be suggested. The regression equations merely indicated the existence of a regional differential for women. They did not indicate the sources of the differential.

b. Suggestions for Further Research. First, and most important, studies on interarea differentials should be developed with more current data both to challenge the validity of the results here which were based on 1960 data and to note any changes. Furthermore, more refined measurement of the variables is needed, particularly for such things as the capital-labor ratio. A direct measure of earnings rather than an estimate can offer more precision.

---

[12] Evening Newscast, Channel 2, Chicago, Illinois, 11 September 1974.

Above all, the study indicates a need for exploration into the forces that determine the earnings of women. It has been the custom to treat both men and women symmetrically in wage theory, but the forces which determine the earnings of women seem to be clearly different from those which determine the earnings of men.

BIBLIOGRAPHY

BIBLIOGRAPHY

Becker, Gary S. The Economics of Discrimination. Chicago: University of Chicago Press, 1957.

Becker, Gary S. Human Capital. New York: National Bureau of Economic Research, 1964.

Becker, Gary and Barry R. Chiswick. "Education and the Distribution of Earnings," Papers and Proceedings, American Economic Review, 56, May 1966, pp. 358-392.

Bell, Duran. "Occupational Discrimination as a Source of Income Differences: Lessons of the 1960's," American Economic Review, 62, May 1972, pp. 363-372.

Bloch, Joseph W. "Regional Wage Differentials, 1907-1946," Monthly Labor Review, 66, April 1948, pp. 371-377.

Bronfenbrenner, Martin. Income Distribution Theory. Aldine-Atherton, Inc., Chicago, 1971.

Cain, Glen, W. Lee Hansen, and Burton A. Weisbrod. "Occupational Classification: An Economic Approach," Monthly Labor Review, 90, February 1967, pp. 48-52.

Christ, Carl. Econometric Models and Methods. John Wiley & Sons, Inc., New York, 1966.

Coelho, Philip R. P. and Moheb A. Ghali. "The End of the North-South Wage Differential," American Economic Review, 61, December 1971, pp. 932-937.

Coelho, Philip R. P. and Moheb A. Ghali. "The End of the North-South Wage Differential: Reply," American Economic Review, 63, September 1973, pp. 757-762.

Douglas, Paul H. The Theory of Wages. New York: The Macmillan Company, 1934.

Douglas, Paul H. "Comments on the Cobb-Douglas Production Function," in The Theory and Empirical Analysis of Production, Studies in Income and Wealth, vol. 31. New York: Columbia University Press, 1967.

Douty, Harry M. "Wage Differentials: Forces and Counterforces," Monthly Labor Review, 91, March 1968, pp. 74-81.

Draper, Norman R. and Harry Smith. Applied Regression Analysis. John Wiley & Sons, Inc., New York, 1966.

Duncan, Otis. "Occupational Components of Educational Differences in Income," Journal of the American Statistical Association, 56, December 1961, pp. 776-788.

Eckstein, Otto and Thomas A. Wilson. "The Determination of Money Wages in American Industry," Quarterly Journal of Economics. Reprinted in Perspectives on Wage Determination, Campbell R. McConnell, ed. New York: McGraw-Hill Book Co., 1970.

Ferguson, C. E. The Neoclassical Theory of Production and Distribution. Cambridge University Press, 1969.

Fisher, Franklin M. "Approximate Aggregation and the Leontief Conditions," Econometrica, 37, March 1969, pp. 457-469.

Fisher, Franklin M. "The Existence of Aggregate Production Functions," Econometrica, 37, October 1969, pp. 553-557.

Fisher, Franklin M. "Aggregate Production Functions and the Explanation of Wages: A Simulation Experiment," Review of Economics and Statistics, 53, November 1971, pp. 305-325.

Fuchs, Victor R. Differentials in Hourly Earnings by Region and City Size, 1959. National Bureau of Economic Research Occasional Paper 101. New York: Columbia University Press, 1967.

Gallaway, Lowell E. "The North-South Wage Differential," Review of Economics and Statistics, 45, August 1963, pp. 264-272.

Griliches, Zvi. "Notes on the Role of Education in Production Functions and Growth Accounting," in Education, Income, and Human Capital, W. Lee Hansen, ed. New York: Columbia University Press, 1970.

Gwartney, James D. and Kenneth M. McCaffree. "Variance in Discrimination Among Occupations," Southern Economic Journal, 38, October 1971, pp. 141-155.

Hall, Robert E. and Richard A. Kasten. "The Relative Occupational Success of Blacks and Whites," Brookings Papers on Economic Activity, 3, 1973, pp. 781-797.

Hicks, John R. The Theory of Wages. Second Edition, New York: St. Martin's Press, Inc., 1963.

Houff, James N. "Area Wages and Living Costs," Monthly Labor Review, 92, March 1969, pp. 43-46.

Johnson, George E. "The Demand for Labor by Educational Category," Southern Economic Journal, 37, October 1970, pp. 190-204.

Johnson, Harry G. and Peter Mieszkowski. "The Effects of Unionization on the Distribution of Income: A General Equilibrium Approach," Quarterly Journal of Economics, 84, November 1970, pp. 539-561.

Kassalow, Everett M. "What Happens When Everyone Organizes?" Monthly Labor Review, 95, April 1972, pp. 27-32.

Ladenson, Mark L. "The End of the North-South Wage Differential: Comment," American Economic Review, 63, September 1973, pp. 754-756.

Leiserson, Mark W. "Wage Decisions and Wage Structures in the United States," in Wage Structure in Theory and Practice, E. M. Hugh-Jones, ed. Amsterdam: North-Holland Publishing Co., 1966.

Lester, Richard A. "Diversity in North-South Wage Differentials," Southern Economic Journal, 13, January 1946, ppl 238-262.

Lester, Richard A. "A Range Theory of Wage Differentials," Industrial and Labor Relations Review, 5, July 1952, pp. 483-500.

Lydall, Harold. The Structure of Earnings. Oxford: The Clarendon Press, 1968.

Lewis, H. Gregg. Unionism and Relative Wages in the United States. University of Chicago Press, 1963.

Machlup, Fritz. "Theories of the Firm: Marginalist, Behavioral, Managerial," American Economic Review, 57, March 1967, pp. 1-33.

Mansfield, Edwin. "City Size and Income, 1949," in Regional Income, National Bureau of Economic Research, 1957.

Mayhew, Anne. "Education, Occupation, and Earnings," Industrial and Labor Relations Review, 24, January 1971, pp. 216-225.

Miller, Herman P. Income of the American People. New York: Wiley Publishing Co., 1955.

Miller, Herman P. Income Distribution in the United States, (1960 Census Monograph), U.S. Government Printing Office, 1966.

Mincer, Jacob. "The Distribution of Labor Incomes: A Survey with Special Reference to the Human Capital Approach," Journal of Economic Literature, 8, March 1970, pp. 1-26.

Nadiri, M. Ishaq and S. Rosen. "Interrelated Factor Demand Functions," American Economic Review, 59, September 1969, pp. 457-471.

Nadiri, M. Ishaq. "Some Approaches to the Theory and Measurement of Total Factor Productivity: A Survey," Journal of Economic Literature, 8, December 1970, pp. 1137-1176.

Ozanne, Robert. Wages in Practice and Theory. University of Wisconsin Press, 1968.

Rao, Potluri and LeRoy Miller. Applied Econometrics, Wadsworth Publishers, Belmont, California, 1971.

Reder, Melvin. "A Theory of Occupational Wage Differentials," American Economic Review, 45, December 1955, pp. 833-852.

Scoville, James G. "Education and Training Requirements for Occupations," Review of Economics and Statistics, 48, August 1966, pp. 387-394.

Scoville, James G. The Job Content of the U.S. Economy, 1940-1970. New York: McGraw-Hill Book Co., 1969.

Scully, Gerald W. "Interstate Wage Differentials: A Cross-Section Analysis," American Economic Review, 59, December 1969, pp. 757-773.

Segal, Martin. "Regional Wage Differentials in Manufacturing in the Postwar Period," Review of Economics and Statistics, 43, May 1961, pp. 248-255.

Walters, A. A. "Production and Cost Functions: An Econometric Survey," Econometrica, 31, January-April 1963, pp. 1-66.

Weiss, Leonard W. "Concentration and Labor Earnings," American Economic Review, 56, March 1966, pp. 96-117.

Welch, Finis. "Education in Production," Journal of Political Economy, 78, January/February 1970, pp. 1-18.

Welch, Finis. "Labor Market Discrimination: An Interpretation of Income Differences in the Rural South," Journal of Political Economy, 75, June 1967, pp. 225-240.

Welch, Finis. "Black-White Differences in Returns to Schooling," American Economic Review, 68, December 1973, pp. 893-907.

APPENDIX

TABLE A1

AREAS STUDIED IN THE EMPIRICAL ANALYSIS

* Birmingham, Alabama	Albany-Schenectady, New York
Phoenix, Arizona	Buffalo, New York
L.A.-Long Beach, Cal.	New York, New York
San Bern.-Riv-Oak, Cal.	Akron, Ohio
San Fran.-Oakland, Cal.	Canton, Ohio
* Wilmington, Delaware	Cincinnati, Ohio
* Miami, Florida	Cleveland, Ohio
* Atlanta, Georgia	Columbus, Ohio
Chicago, Illinois	Dayton, Ohio
Davenport-RI-Moline, Ill.	Toledo, Ohio
Wichita, Kansas	Portland, Oregon
* Louisville, Kentucky	Allentown, Pennsylvania
* New Orleans, Louisiana	Philadelphia, Pennsylvania
* Baltimore, Maryland	Pittsburgh, Pennsylvania
Boston, Massachusetts	Providence, Rhode Island
Worcester, Massachusetts	* Chattanooga, Tennessee
Detroit, Michigan	* Memphis, Tennessee
Minnesota-St. Paul, Minn.	* Dallas, Texas
Kansas City, Missouri	* Fort Worth, Texas
St. Louis, Missouri	* Richmond, Virginia
Jersey City, New Jersey	Seattle, Washington
Newark, New Jersey	Milwaukee, Wisconsin
Paterson-Clifton, N.J.	

* South, as classified by the Bureau of the Census
All other areas were classifed as North