Labor mobility (1981 folder)

Intergenerational Occupational Mobility in the United States

A Segmentation Perspective

Marshall I. Pomer

University of Florida Social Sciences Monograph Number 66

Labor mobility
(1981 folder)

Intergenerational Occupational Mobility in the United States:

A Segmentation Perspective

Marshall I. Pomer

University of Florida Monographs:
Social Sciences Number 66

A University of Florida Book

University Presses of Florida [1981]
Gainesville

University Presses of Florida is the central agency for scholarly publishing of the State of Florida's university system. Its offices are located at 15 NW 15th Street, Gainesville, FL 32603. Works published by University Presses of Florida are evaluated and selected for publication by a faculty editorial committee of any one of Florida's nine public universities: Florida A&M University (Tallahassee), Florida Atlantic University (Boca Raton), Florida International University (Miami), Florida State University (Tallahassee), University of Central Florida (Orlando), University of Florida (Gainesville), University of North Florida (Jacksonville), University of South Florida (Tampa), University of West Florida (Pensacola).

Library of Congress Cataloging in Publication Data

Pomer, Marshall.
Intergenerational occupational mobility in the United States.

(University of Florida monographs: Social sciences; no. 66)
"A University of Florida book."
Bibliography: p.
Includes index.
1. Occupational mobility--United States. 2. Social mobility--United States. I. Title. II. Series: Florida. University, Gainesville. University of Florida monographs: Social sciences; no. 66.
HD5717.5.U6P65 305.5′0973 80-20626
ISBN 0-8130-0674-0

PRINTED IN U.S.A.

CONTENTS

LIST OF TABLES

ACKNOWLEDGMENTS

I AM GRATEFUL to several of my teachers for their influence: Zvi Griliches inspired me to strive for cautious and rigorous data analysis; Stephen Marglin showed me the ideological bias in economic theory; Sam Bowles and Herbert Gintis convinced me to pursue the notion of social class; Christopher Jencks stirred my interest in measuring social mobility.

In addition, I wish to thank many other colleagues for their interest in this study and for their helpful comments, especially Robert Alford, Paul Attewell, Ronald Breiger, Carolyn Clark, William Domhoff, William Frazer, Marcia Freedman, Benjamin Gorman, Bennett Harrison, John Henretta, James Henry, G. S. Maddala, Peter Meyer, and Lee Rainwater.

This study is limited to data on white males. I apologize to those who are struggling against sexist and racist tendencies in our society and in social research. The restriction in focus reflects my intent to reexamine the same data which have been widely used to rationalize the neglect of class in social research.

Thanks must go also to the Graduate School of the University of Florida for making possible the publication of this monograph.

To
JEREMY and JOSHUA

CHAPTER ONE

INDIVIDUALISTIC ANALYSIS

ACCORDING to the American Dream, you can be anything you want as long as you have the desire and the ability. A long-standing function of sociologists has been to show that the reality is not the same as the ideal, that life chances are severely affected by socioeconomic origins. More recently, however, quantitative analysis has tended to create the impression that the United States is close to its ideal.

The shift toward quantitative methodology and optimistic assessment of mobility began near the height of the War on Poverty with the work of Blau and Duncan (1965, 1967). They revolutionized the study of mobility by assembling a large national sample and by introducing novel statistical techniques. Their data and methods led to the impression, most strongly evident in the work of Jencks (1972), that socioeconomic origins are not very consequential.

Blau and Duncan's data, which are reanalyzed in this study, are from the Survey of Occupational Changes in a Generation, or more briefly, the OCG sample. The OCG sample consists of all men aged 20 to 64 in the civilian, noninstitutional population who were selected for the March 1962 Current Population Survey by the U. S. Census Bureau and who returned a supplementary questionnaire. The questionnaire included questions on occupation and education of father when respondent was 16 years old. The OCG sample is large (20,700 men) and designed to be representative of the U. S. as a whole. It is regarded as "the first comprehensively designed inquiry into social mobility in the United States" (Heller, 1969:313). Until recently, it was the largest available data base on intergenerational mobility (Featherman and Hauser, 1978).

Blau and Duncan, hesitant to state directly that socioeconomic origins are not important, indirectly downgrade the role of origins. For example, they (1967:198-200) criticize the conclusion of a well-known study of a metropolitan labor market that "occupational and social status are to an important extent self-perpetuating" (Lipset and Bendix, 1959:198). And Duncan (1969:109) supports another researcher's recommendation, based on the OCG sample, to reject "the idea that the economic status of parents is a major determinant of the economic status of offspring" (Gallaway, 1966:7). Duncan, however, is direct in his treatment of "inheritance of poverty," terming the "cycle of poverty" a myth (Duncan, 1969: especially 85-87; also Blau and Duncan, 1967:199-205).

Jencks (1972), in a widely read analysis of inequality, is still more outspoken. Using the methods and data of Blau and Duncan, Jencks states, "The role of a father's family background in determining his son's status is surprisingly small, at least compared to most people's preconceptions" (p. 179). And there is "an enormous amount of economic mobility from one generation to the next" (Jencks, 1972:7). In addition to general assertions downplaying the role of socioeconomic origins, Jencks (1972:7) specifically attacks the belief that socioeconomic origin is basic to the analysis of poverty, calling it "erroneous" to assume that "middle-class children rarely end up poor."

In sum, the judgment that origins are unimportant can be broken down into four claims:

(1) There is a large amount of intergenerational mobility.
(2) The effect of socioeconomic origins on life chances is relatively minor.
(3) Poverty is not passed down from one generation to the next.
(4) Elimination of the effect of origins on life chances would have only a minor impact on overall socioeconomic inequality.

One might expect separate evidence for each of these four claims. This is not the case. Instead, the claims are loosely tied to a questionable interpretation of a crude statistic. As we shall explain, the correlation between the socioeconomic positions of fathers and sons is misinterpreted.

A person's socioeconomic position is crudely measured by the socioeconomic index (SEI) value associated with his or her census occupation. The SEI value for each detailed census occupation is determined by the educational and income characteristics of males in that occupation (for details, see Duncan, 1961; and Blau and Duncan, 1967:117-28). The socioeconomic position of respondent's father is measured by the SEI score of his occupation when the respondent was age 16 (as reported by respondent at time of the survey).[1] The socioeconomic position of the respondent is measured by the SEI score of respondent's occupation at the time of the survey.

Once the socioeconomic positions of fathers and sons are measured, a father-son SEI correlation coefficient is easily calculated. A hazardous next step is to give a reduction-in-inequality interpretation to the calculated value for the correlation coefficient.

The reduction-in-inequality interpretation compares inequality in SEI among all men to inequality in SEI among men with the same socioeconomic origin. The reduction-in-inequality percentage is the percentage amount by which inequality among men with the same origin is less than inequality among all men. An important technical requirement for defining this concept is that the amount of inequality among men with the same origin is the same for every origin.

By making a strong assumption about the form of the relationship between father's and son's socioeconomic position, the reduction-in-inequality percentage may be deduced from the father-son SEI correlation coefficient. The required assumption is that father's and son's SEIs are random variables which conform to the bivariate normal probability distribution. No evidence is offered in support of this assumption.[2]

The reduction-in-inequality interpretation requires that a choice be made as to how to define inequality mathematically. Jencks uses the standard

1. The Duncan SEI is only one of many measures of socioeconomic position. Limited by the available data, social stratification researchers from a variety of theoretical perspectives have relied heavily on SEI to measure socioeconomic position of the individual. Income is an alternative measure which may be more comfortable for economists. Regardless of the choice of measure of socioeconomic position, the same problems arise as to how to interpret the father-son correlation.

2. If joint-normality is not assumed, the reduction-in-inequality percentage cannot be interpreted as a comparison of overall and subgroup inequality. Instead, this percentage could be viewed as a technical indicator of the goodness of fit of a linear regression. This more cautious approach is more consistent with Blau and Duncan (1967) than Jencks (1972), and is less likely to lead to the hasty conclusion that socioeconomic origins are unimportant.

deviation or multiples of the standard deviation. Before criticizing the choice of the standard deviation, let us consider the estimates for reduction-in-inequality obtained by Jencks and by Blau and Duncan, and the revisions to these estimates obtained by adjusting for measurement error.

Applying the reduction-in-inequality method to a computed SEI correlation of 0.38, Jencks finds that inequality among men with the same socioeconomic origin is only 7.5 percent less than inequality among all men. Thus he concludes that "there is nearly as much inequality among men whose parents had the same economic status as among men in general" (Jencks, 1972:8; see also 179, 215, 220). The same argument implies that origins are not important for understanding income inequality (Jencks, 1972:chap. 7). The impression is also left that taking wealth into account will not alter the general conclusion that origins are unimportant (Jencks, 1972:214).

Blau and Duncan (1967) use the variance rather than the standard deviation to measure inequality in SEI. If bivariate normality is assumed as above, the SEI correlation may be interpreted as meaning that inequality among men having the same socioeconomic origins is 85.6 percent as great as the inequality among all men. "The overall correlation between father's and son's occupational status is 0.38. This indicates that there is much occupational mobility in the United States; only one-seventh of the variance in economic status is attributed to the influence of father's socioeconomic status" (Blau and Duncan, 1965:6).

Blau and Duncan consider several variants of this approach. Socioeconomic origins may be classified by interval of SEI (multiple-classification analysis). Using regression analysis, additional variables may be introduced to measure social origins (in particular, father's education). In general, the conclusion is that inequality for persons with identical origins is about 85 percent of overall inequality.

The amount of inequality reduction may be revised considerably upward by allowing for measurement error. Fairly conservative assumptions suggest that the father-son SEI correlation may be revised upward to 0.50. (Correcting for measurement error, Jencks [1972:200 n.13] estimates the correlation to be 0.483.) An increase in the SEI correlation from 0.38 to 0.50 implies substantial increases in the reduction-in-inequality estimates. The estimated 7.5 percent reduction in inequality based on the standard deviation is increased to 13.4 percent. The 14.4 percent reduction based on the variance is increased to 25.0 percent.

A reduction in inequality of 25 percent suggests that 25 percent of inequality is due to socioeconomic origins and that the remaining 75 percent is due to other factors. Similarly, when it is stated that "only one-seventh of the variance in socioeconomic status is attributable to the influence of father's socioeconomic status" (Blau and Duncan, 1965:6), it is implicit that six-sevenths of the variance is due to other factors besides level of socioeconomic origin. Such an interpretation requires that the variance and not the standard deviation be used, and that a number of assumptions be made.[3] A very simplistic model of intergenerational mobility is assumed to be correct–a model which in turn is founded on simplistic assumptions about the economy and about social stratification. Technically, the partitioning of variance depends on the assumption of a linear, additive model.[4] For example, such a model ignores that the impact of socioeconomic origins is affected by regional differences in educational and employment opportunity.

The American Dream suggests that variation in socioeconomic position is not due to unequal opportunity but to differences in effort and ability. Therefore, variation in socioeconomic position not attributable to socioeconomic origins tends to be equated with equal opportunity. Thus, Blau and Duncan attribute the unexplained variation in SEI to equal opportunity in their attack on the notion of a poverty cycle (1967:199-205). Jencks frequently attributes unexplained variation to luck, which is different from ability and effort but is consistent with an equal opportunity view of the labor market (Jencks, 1972:8, 9, 131, 227, 228).

A number of types of unequal opportunity are not captured by level of socioeconomic origin. There are forces of prejudice and discrimination. Some of the effect of family on aspirations and social connections is independent of socioeconomic position. Educational opportunity is not the same for all persons with the same socioeconomic position due to differences in the quality of public schools and the availability of scholarships. There is also variation over time and place in employment opportunity, which is consequential for both launching and maintaining a career.

Employment opportunity is affected by macro changes in the supply of jobs. Differences in the organizational structure of occupations and enterprises, and differences in their stability and growth, are also consequential for job security and advancement opportunity. Thus individuals face different life chances depending not only on their locality and the time at which

3. Let X and Z be uncorrelated and $Y = aX + bZ$. Then $V(Y) = V(Y|X) + V(Y|Z)$, but $STD(Y) < STD(Y|X) + STD(Y|Z)$.

4. Let X and Z be uncorrelated and $Y = aX + bZ + cXZ$. Then $V(Y) < V(Y|X) + V(Y|Z)$.

they enter the labor market but also on the characteristics of the occupation or enterprise in which they gain "establishment" (Freedman, 1969).

An alternative approach to assessing strength of relationship is direct examination of the regression coefficient implied by the SEI correlation. Assuming an 0.5 correlation coefficient and adjusting for the fact that the variance in SEI is almost 40 percent greater for sons than for fathers,[5] the slope of the regression line is about 0.6. Accepting this figure, one summary interpretation is that about 60 percent of the advantage in socioeconomic standing of the father is passed on to his son. Yet, as we have discussed, the same correlation implies that the standard deviation of SEI controlling for SEI of origin is only 13.4 percent less than the overall standard deviation. That is, although two individuals having father's SEI ten units above the average have an expected difference in SEI that is 87 percent of the expected difference for two individuals drawn at random, they both have an expected SEI six units above the average. Thus the value of the correlation between the socioeconomic positions of father and son need not be interpreted as an indication of a weak relationship. Very little of the recent quantitative research takes issue with the view that the effect of origins is not large. A major exception is the work of Bowles and Gintis (Bowles and Gintis, 1972; Bowles, 1972), which relies heavily on conjectures concerning measurement error. Brittain (1977) has also rebutted the thesis that origins are unimportant, but he does so with a very small and possibly ungeneralizable sample.

It appears that Duncan may have minimized the force of inheritance in his concern to show that the low socioeconomic positions of blacks cannot be accounted for by their low socioeconomic origins (in particular, Duncan, 1969). Jencks may have downplayed the role of origins due to his wish to go beyond equality of opportunity to equality of condition for the successful and the unsuccessful alike (Jencks, 1972:chap. 9). The widespread impression that socioeconomic origins are unimportant may thus be due to attempts to focus attention on other factors.

Schooling is the factor that has received the most attention. Much of the assessment that origins are not very important rests on the empirical result that, once the number of years of schooling is accounted for, the effect of origins is greatly reduced and is small in comparison to the effect of years of schooling.

Nevertheless, the gross connection between origins and destinations is itself of central interest in characterizing the equality of a society. Account-

5. See chapter 5.

ing for the connection is a separate task, and one which is likely to be misleading unless the gross connection is assessed correctly.

The remainder of this introduction outlines chapter by chapter the steps taken to reanalyze the OCG data on intergenerational mobility. Featherman and Hauser (1978) have published a study of a replicate of the OCG survey. Unfortunately their study is too recent to permit analysis of the replicate data here or to appraise the results and conclusions. They appear, however, to be at least as optimistic as Blau and Duncan in their conclusion that opportunity has increased since the earlier Blau and Duncan study.

Chapter 2 discusses conceptual approaches to studying labor and develops the distinction between individualistic and segmentation views. A variety of traditions that have rejected individualistic analysis are reviewed as examples of the segmentation perspective.

The discussion of labor market theory in chapter 2 emphasizes that strong assumptions are required in order to make policy predictions on the basis of the cross-sectional relationship between origins and destinations. Based on the result that holding origins constant reduces the standard deviation of socioeconomic position by only about 10 percent, Jencks concludes that there is little value in policies that reduce the effect of origins on life chances. Apart from doubting the accuracy of the estimates and the use of the standard deviation to measure socioeconomic inequality, one can question what appears to be a naive application of a competitive model of the labor market. Properly applied, a competitive model implies that a policy such as upgrading the schooling of those with low origins would have a leveling effect that is not captured by the cross-sectional data: not only should such a policy increase the marginal productivity (and therefore the earnings) of those with low origins; it should also, by increasing the supply of human capital, reduce the marginal productivity of the well-schooled, who are disproportionately of high socioeconomic origins. On the other hand, a segmentation model, emphasizing the power of workers organized in professions and unions to restrict entry and to strike, would imply that changing the characteristics of workers would have little effect on distributional inequality—unless those changes affected worker organization.

Chapter 3 examines the effect of origins by comparing destination probabilities for persons of different socioeconomic origins. Even a fairly low father-son correlation on SEI implies, under the assumption of joint normality, that the chances of ending up at or near the top (or the bottom) of the socioeconomic structure are very unequally distributed according to socioeconomic origin (Bowles and Gintis, 1972; Bowles, 1972). Chapter 3, however, avoids the assumption of normality and the Duncan SEI index: it directly calculates the conditional probabilities from observed frequencies. By comparing by origins the incidence of low socioeconomic position, this chapter provides direct evidence that poverty is, to an important extent, hereditary.

Rather than ask how completely are life chances determined by origins, chapter 4 instead develops socioeconomic categories based on similarity of life chances and similarity of social origins. Chapter 4 uses the mobility data to determine highly aggregative categories via cluster analytical techniques. This chapter may be narrowly viewed as providing a basis for collapsing row and column categories of the mobility table so as to make possible a more parsimonious description of intergenerational mobility. More broadly conceived, the objective is to use mobility data to build a model of socioeconomic structure. In this regard, the categories derived in chapter 4 may be compared to the continuous model of socioeconomic structure and the various segmentation models of labor reviewed in chapter 2.

Thus, chapters 3 and 4 wholly abandon the reduction-in-inequality approach. Instead, these chapters develop and apply methods that use as their starting point a set of empirical probabilities of moving from each socioeconomic origin to each socioeconomic destination. This approach avoids simplistic assumptions concerning the form of the relationship between origins and destinations. In addition, comparison of destination probabilities provides a direct assessment of whether socioeconomic origins strongly affect life chances.

Chapter 5 considers a reduction-in-inequality approach but does so using a model that is less simplistic than the one previously used to interpret the father-son SEI correlation. Labor segments are used to define socioeconomic origins, and the mathematical form of the effect of origins is not assumed to be the same for all origins. Regression analysis is applied with an emphasis on the use of robust techniques.

In sum, previous studies have incautiously interpreted the father-son correlation in socioeconomic position. They have used a reduction-in-inequality interpretation of the observed correlation to downgrade the importance of socioeconomic origins. We have discussed numerous difficulties with this approach, including (1) the inaccuracy of the estimated correlation due to the crudeness of the model and the data; (2) the sensitivity of the interpretation to the mathematical definition of inequality; (3) the question of what amount of reduction in inequality is to be regarded as substantial; (4) ambiguity concerning unmeasured influences and the interaction between socioeconomic origins and the unmeasured influences. Chapter 2 argues for a segmentation perspective in place of the individualistic conceptualization which underlies the reliance on the father-son correlation in socioeconomic position. Chapters 3, 4, and 5 apply the segmentation perspective.

CHAPTER TWO

LABOR SEGMENTS

THE RESEARCH described in the previous chapter is individualistic. Inequality is understood as the consequence of differences among individuals. An alternative is to use the group rather than the individual as the basic unit of analysis. From this more sociological perspective, inequality is described by comparing groups rather than individuals, and inequality is explained by conflict among groups rather than by competition among individuals.

Individualistic analysis dominates orthodox economics. There are, however, alternative approaches in labor economics which use some notion of 'group' or 'labor segment' as the basic unit of analysis. This chapter describes several examples of segmental analysis and defends the segmentation approach against criticisms arising from orthodox economics.

Segmentation approaches

The various segmentation approaches differ in their definitions of labor segments. For example, a prescribed segment might be self-employed professionals, unskilled black males, or the United Farm Workers. It is sometimes unclear as to whether labor segments are to be thought of as a set of jobs or a set of workers. Although it is most useful to regard a labor segment as, by definition, a set of jobs, labor segments are almost invariably defined in such a way that workers of a particular labor segment constitute a meaningful social group.

We shall examine three traditions in labor economics that embody the segmentation approach. The theory of "noncompeting groups," current in the second half of the nineteenth and early twentieth centuries, conceived of labor segments as hereditary social classes. The "dual market" model, popular since 1970, characterizes the labor market as consisting of inferior jobs held by women and blacks and superior jobs largely monopolized by white men. The "institutional view," which has commanded attention in varying amounts throughout most of the twentieth century, emphasizes bargaining units such as unions.

These three approaches to labor segmentation differ in the degree to which they aggregate the basic unit of analysis. The bargaining units of institutional analysis may be regarded as components of the broad categories of either the dual view or the noncompeting groups model. The dual view puts institutional analysis in the perspective of a society divided by race and sex; the noncompeting groups model puts institutional analysis in the perspective of a society divided according to differences in culture and economic resources.

Noncompeting groups.– As formulated by Taussig (1921), there are five labor segments. The bottom segment consists of the "day laborers." The second consists of the "typical manual laborers," who enjoy greater continuity of employment than day laborers but lack trade union protection and the "trained hand" of the "skilled workmen." Skilled workmen, the "aristocracy of the manual laboring class," make up the third segment. The fourth is the "lower middle class" of "clerical or semi-intellectual" workers. The most privileged segment is the "class of the well-to-do" made up of the "professionals," "salaried officials," and "business men and managers of industry."

According to the noncompeting-groups approach, schooling and values are important for transmitting segment membership between generations. Children of the day laborers rarely carry their education beyond the minimum because of family pressures for the children to begin work. Skilled workers have "pride in occupation" and are well able to support their children through grammar school and apprenticeship. The lower-middle class and the well-to-do have a "contempt for manual work," but the lower-middle class has less access to elite schooling.

The labor segments define a small number of broad strata, which may be regarded as social classes. In addition to the tendency toward hereditary membership and different values for each segment, the segments define boundaries for social interaction and have internal homogeneity. On the other hand, as suggested by the term "noncompeting," the labor segments are not class conscious in the Marxian sense of being agents of history that struggle for the interests of their members.

Dual labor market.—The popularization of the dual labor market model beginning about 1970 brought a resurgence of interest in the segmented view of labor (see Piore, 1969, 1970, 1975; Bluestone, 1970; Doeringer and Piore, 1971; Harrison, 1972a, 1972b; Gordon, 1971, 1972; O'Connor, 1973; Edwards et al., 1975). The dual model divides jobs in the economy into primary and secondary jobs. Secondary jobs are characterized by low pay and poor working conditions. Primary jobs are better paid, have better working conditions, and offer more job security and more favorable chances of advancement.

Primary jobs, especially for manual workers, are found disproportionately in large firms that sell in the less competitive markets. The economic and political power of the large firms, often reinforced by large unions, weakens the smaller firms and inhibits the creation of primary jobs in smaller firms.

Whether it be nonunion IBM or strongly unionized General Electric, workers in primary jobs have job retention rights and are often assured pay increases and a chance for promotion if they remain loyal to their jobs. These job characteristics are attributed to an elaborate hierarchical job structure with hiring confined to the first rung of job ladders. Marxist economists have attributed the extreme degree of hierarchy to the need to divide and control workers (Braverman, 1974; Marglin, 1975; Edwards et al., 1975). This explanation is consistent with Kalecki's (1970) stress on the problem of protecting monopoly profits from organized labor, but it is in conflict with the emphasis by neoclassical economists on human capital formation as the result of on-the-job training (Becker, 1964).

According to the dual labor market approach, most primary jobs are accessed by normal progression through the internal hierarchy of a firm. Access to the entry level primary jobs is determined by a rationing process rather than a market equilibrating process. Because of the force of prejudice and social networks, women and blacks are disproportionately excluded. The policy implication of the dual labor market model is to improve the employment opportunities for minorities.

Researchers interested in sexual and racial inequality invariably group workers by race or sex or both. In the dual labor market model, two labor segments are defined as the primary workers (consisting of prime age white males) and the secondary workers (consisting of teenagers, women, and blacks). Defining labor segments on the basis of such ascriptive traits reinforces cultural stereotypes and fails to focus attention on access to primary jobs and the creation of primary jobs. Labor segment mobility cannot be defined, of course, if the labor segments are defined solely by ascriptive traits.

In contrast to orthodox economics, the dual approach pays considerable attention to industrial structure. Also unlike orthodox economics, the dual approach recognizes excess supply of qualified workers and bureaucratization of workers.

Compared to advocates of the noncompeting-groups approach, advocates of the dual labor market approach have tended to neglect social class formation based on other than racial differences.[1] Apart from racial and sexual discrimination, inheritance of position in the labor market is seldom mentioned. As in the case of the noncompeting-groups approach, little measurement of the size of labor segments is carried out. Also the two segments are not regarded as cohesive or capable of acting as agents. Thus the dual approach has directed attention away from worker organization, which has been a more central concern for the institutional approach.

Institutional approach. – Institutional labor market analysis has been concerned largely with the problem of industrial relations. For much of this century, labor unrest and the growth of unionization were major public issues, and they constituted the special domain of labor economists. As a result, institutional labor economists traditionally focused almost entirely on workers who were part of the unionized sector, primarily blue-collar workers. The dual labor market model as originally developed by Piore and Doeringer, and other labor economists, reflects this same limitation of focus (though seldom explicitly).

A new version of the institutional approach has given attention to the full range of occupations while maintaining the focus on worker organization. The "sheltering theory" of the labor market is concerned with the construction or "establishment" of shelters by the organized effort of workers or due to the threat of such organized effort, and in reflection of the structure of product markets. Freedman (1976), who coined the term, has classified jobs in terms of degree of shelter.

A major theme of the institutional approach is the workers' ability to restrict entry into their occupations or "shelters." However, there is little discussion of the effect that workers' control over entry into their occupations has on occupational inheritance.

Insofar as an occupation consists of similarly sheltered workers, it makes sense from this perspective to regard occupations as the basic units of analy-

1. An exception is Piore (1975).

sis. It is often necessary, however, also to consider segmentation by industry or firm of the members of an occupation; and it is necessary to recognize the heterogeneity of broadly defined occupations.

For the institutional approach, a group of workers who operate to some degree as a cohesive unit is regarded as the basic unit of analysis. These labor segments may be viewed as components of the broad segments that are defined by either the noncompeting groups model or the dual labor market model.

Other applications.– The three traditions in labor economics discussed above had to contend with attack from orthodox economics, which will be considered in the next section. First, however, we shall consider other examples of research that can be considered to reflect a segmentation perspective. These examples are not integral parts of the segmentation traditions just discussed and were not self-consciously developed in opposition to orthodox economics. In most of these examples, a labor segment is an occupation.

Labor economists have contributed to the analysis of trends in income distribution by examining cyclical and secular changes in occupational wage differentials (for example, Keat, 1960). A neglected question is whether the occupational categories used represent the most interesting definition. Some of these studies explain the changes in occupational wage structure on the basis of labor quality (for example, Reder, 1955), others emphasize demand (for example, Okun, 1973), and some emphasize organization of occupational groups (for example, Maher, 1961). Study of the cyclical variation in occupational wage structure represents a disaggregated analysis of the Phillips curve (Jackson and Jones, 1973); such study has become a pertinent and controversial area of research due to simultaneous inflation and unemployment. The cyclical analysis of industrial wage differentials rather than occupational wage differentials (as in Wachter, 1970) may be thought of as based on an industrial rather than occupational definition of labor segments. Okun's (1973) analysis of the distributional consequences of macroeconomic policy represents an effective use of the dual labor market model.[2] Okun shows the beneficial effect of low unemployment on the representation of blacks and women in primary jobs. He also points out that the relative pay for secondary jobs rises when unemployment is reduced. Okun, however, leaves the impression, in contradiction to the noncompeting-groups view, that only blacks, women, and teenagers have secondary jobs—that low

2. Typically studies of the distributional effects of economic policy disregard occupational structure: for example, Gramlich (1974) and Smith et al. (1974).

socioeconomic origins do not significantly influence the occupational prospects of white males. The institutional view points to other factors neglected by Okun, including control over licensure and selection for apprenticeship, and the long-term effect of low unemployment on the strength of worker organization.

The sociology of occupations also embodies a segmentation perspective. Sociologists in this field pick narrowly defined specific occupations as subjects of investigation. Specific professions (such as medicine) or types of work (such as lumberjacking) are studied as determinants of social identity, self-conception, and social status. Sociologists have also investigated how occupations develop and maintain power. (For a review, see Hughes, 1959.) The sociology of occupations has begun to be influenced by the segmentation approaches employed in labor economics (for example, see Montagna, 1977). Sociologists investigating social stratification also are concerned with defining labor segments (see Wright and Perrone, 1976).

Orthodox criticism

Orthodox economists respond to the segmentation approach with two important claims.[3] First is that the main difference between the segmentation approach and orthodox economics is in regard to degree of competition. Second is that the segmentation approach is theoretically inferior.

The remainder of this chapter develops five points in defense of the segmentation approach against the orthodox view:

(1) The main difference between the segmentation approach and orthodox economics pertains to whether the individual or the group is the basic unit of analysis.

(2) The charge that the segmentation approach is theoretically inferior is based on the questionable assumption that orthodox economics is itself successful.

(3) The segmentation approach, regardless of assumptions as to degree of competition, directs attention toward the job structure, which is neglected by orthodox economics.

(4) The assumptions of orthodox economics make it difficult to study unequal opportunity.

3. Since the resurgence of interest in the segmentation approach about ten years ago, two extended reviews of the segmentation approach have been published by orthodox economists: Wachter (1974) and Cain (1976). In addition, Mincer (1970) criticized the noncompeting groups approach in his review of labor economics.

(5) Orthodox economics and the segmentation approach differ in ideological bias.

The main difference.– As noted by Mincer, classical economists took two approaches to inequality in earnings. The first is the principle of "compensatory differentials" or the "competitive hypothesis" first enunciated by Adam Smith, and the second is the principle of "noncompeting groups" advanced by Cairnes and John Stuart Mill. Mincer contrasts the two approaches:

> Smith's compensatory principle is conditional on the strength of competitive forces in the labor market. Labor mobility produces earnings differentials which tend to equalize "net advantages and disadvantages" of work. Second, Mill's and Cairnes's doctrine of "noncompeting groups" proclaims in effect the absence of labor mobility resulting in real income differences, produced and perpetuated by socially, legally, and culturally imposed and inherited stratifications (Mincer, 1970:3).

Thus to orthodox economists, the main difference between the two approaches is whether competition is acknowledged. Competition, however, does not clearly distinguish orthodox economics from the segmentation approach. Advocates of the segmentation approach acknowledge that there is competition, and orthodox economists admit restrictions to competition. If competition were the distinguishing feature, the two approaches would differ only in degree rather than in kind.

The main difference between them is the unit of analysis used: for orthodox labor economists, the individual worker; for the segmentation approach, the group.

The use of labor segments as the units of analysis is especially compelling when the labor segments constitute meaningful social groups. Carter and Carnoy, who see the unit of analysis as the key difference between the segmentation approach and orthodox economics, define a labor segment as a group of persons sharing a particular labor market situation: "the primary unit of analysis is no longer the individual and his free choices, but rather groups or classes who face objectively different labor market situations" (Carter and Carnoy, 1974:21). The significance of the social groups defined by the labor segments is further enhanced to the degree that the segments define boundaries of social interaction, collective identity, and collective action.

From the orthodox perspective, it is generally assumed all workers compete with each other, and thus it does not make much sense to regard workers as facing different labor market situations. It is also generally assumed that competition among workers goes hand in hand with free and unimpeded mobility, between generations, and within people's careers. All this mobility makes it hard to regard labor segments as defining personal identity or boundaries of social interaction. And it is assumed that workers are entirely concerned with maximizing their own individual advantages, so it is difficult to conceive of labor segments as defining units capable of significant collective action. Furthermore, competition among firms is expected to drive out of business the firms that discriminate in their hiring as well as firms whose workers collectively demand more benefits than they would have received if they had not organized.

The orthodox position is that market competition invalidates the segmentation approach. A better statement of the relation between competition and the segmentation approach is that a high degree of market competition tends to reduce the extent to which labor segments constitute meaningful social groups. The orthodox position may be countered on its own terms by identifying noncompetitive characteristics of markets and by showing evidence of intergenerational and intragenerational immobility. The orthodox position can also be countered by showing that labor segments differ markedly in racial and sexual makeup and in percentages of persons from various socioeconomic origins. Contending, however, on these grounds constitutes quibbling over the degree to which something is true.

More fundamental grounds of contention are differences with regard to what is investigated and what kinds of social policies are considered. Such differences arise as a consequence of the disagreement over whether to choose the individual or the group as the basic unit of analysis.

Inferior theory.– Mincer attacks the noncompeting-groups approach for not producing "any cumulative theoretical developments" (1970:3). The apparent reasons for this attack include his high evaluation of the human capital approach, an unawareness of the history of the segmentation approach, and an inclination toward deductive rigor.

In orthodox economics, human capital theory is the modern approach to studying labor. It emphasizes the elaboration of individual choice of schooling and training. Human capital theory was launched as a discipline in the late 1950s by T. W. Schultz. Schultz (1961) argued that economists, in their excitement over Keynes, had lost sight of the supply side. Schultz suggested that policy should focus more on developing human resources and less on demand. True cumulative theoretical development would have been achieved if human capital theory had led to the perfection of the Keynesian insight into the critical role of demand. Instead, human capital theory has led to an abandonment of the demand side. It has been argued that the main value of human capital theory has been to serve the self-interest of orthodox economists. It has given economists a way to do labor economics in a neoclassical way (Piore, 1973), and a way to do labor economics without paying attention to social structure and the connection of individuals to groups or classes (Bowles and Gintis, 1975; Foley, 1975).

The charge that the segmentation approach has yielded no "cumulative theoretical developments" stems in part from failure to note that students of the labor market since Mill's time have used the segmentation view. Further, the form of segmentation models has developed in accordance with changing social issues. Both Taussig (1921) and Marshall (1947) relied on the notion of noncompeting groups. Studies of unionization and the dual view of the labor market are part of the development of the segmentation perspective.

The units of analysis in orthodox analysis are conceived in such a way that they are mathematically tractable. For example, workers are conceived of as independent maximizers. To advocates of the segmentation approach, the mathematically tractable abstractions of orthodox economics are fatally unrealistic. Orthodox models of labor "proceed from a static, ahistorical framework. . . . In contrast to these orthodox models, the segmentation theories . . . are explicitly historical and focus on historical, systematic forces" (Carter and Carnoy, 1970:21).

The orthodox economist assumes a competitive equilibrium among individuals and tends to ignore institutional and social factors. The advocate of the dual labor market view, on the other hand, assumes that there are essentially two types of jobs. Both approaches are thus based on gross simplification. Just as the advocate of the dual labor market view finds the orthodox model leaves out too much of reality to be acceptable, so does the orthodox economist find the dual labor market model unacceptable. To the nonprofessional economist, however, it is perhaps easier to accept the straightforward abstractions of the dual labor market model than the more rarefied abstractions of the orthodox model. In particular, the dual model is meaningful to the worker who believes that other workers, with capacities similar to his or her own, have jobs that offer more pleasant work, higher pay, more security, and greater advancement potential.

Job structure. – The segmentation approach has advantages even if the labor segments do not constitute meaningful social groups. The act of dividing jobs into groups directs attention toward jobs and away from characteristics of individuals. On a general level, the focus on groups of jobs promotes an understanding of inequality as the consequence of the job structure in an economy rather than as a consequence of characteristics of individuals. On a more specific level, once segments are defined, a number of issues arise, including: (1) In what ways do the jobs in one segment differ from the jobs in other segments? (2) What determines the quantity of jobs (or supply of jobs) in each segment? (3) What controls access to each segment?

An individualistic approach obscures the job structure and issues regarding the supply of, and access to, jobs of differing quality. An illustration is provided by Cain's reaction to the point made by two dual labor market economists that economic distress of minorities stems not from the scarcity of jobs but from the scarcity of primary or "good" jobs, which offer "meaningful employment opportunities" (Doeringer and Piore, 1971:164; quoted by Cain, 1976:1238).

Cain rejects the distinction between a scarcity of jobs and a scarcity of primary jobs. He states, "I do not understand the analytical distinction between a lack of good jobs and a lack of jobs" (1976:1238). Cain explains his point by hypothetically comparing "a laid-off auto worker who is white and thirty-five years old" to "an unemployed, unskilled black worker who is twenty-five years old." Cain says that "the fundamental similarity of the unemployment situation of both types of workers is that the jobs available at the time are not good enough." That is, both hypothetical workers experience a lack of good jobs and hence there is no sense of distinguishing a lack of good jobs and a lack of jobs. He then concludes with a comment regarding the need to look further at "the dissimilarities of the two types of workers."

Cain's difficulty may be attributed to preoccupation with characteristics of workers. To grasp the distinction it is necessary to focus on types of jobs rather than types of people. Rather than compare a white auto worker to an unskilled black, it would be more helpful to compare an unskilled job at a unionized General Motors plant with an unskilled job in a small, nonunion, local laundry. The economic distress of minorities will be little affected by increasing the supply of laundry jobs, but it will be affected if the supply of General Motors jobs is increased. The increased supply of General Motors jobs, however, must be great enough so that its effect is not merely to rehire temporarily laid-off prime age white male auto workers.

Equal opportunity.– One concept of equal opportunity is that workers' productive capacities are fully utilized and rewarded, constrained only by their willingness to work. From a segmentation perspective, it is easy to doubt the existence of equal opportunity since a worker is utilized and paid differently depending on the labor segment he or she occupies. The orthodox model, however, can make it difficult to doubt that opportunity is equal.

The orthodox model holds that competition among workers for jobs, together with competition among firms for workers, yields the happy result that workers identical in ability and effort will be rewarded identically. Restated in terms of the (ideal) market, the market is indifferent to who the seller is, and thus labor of the same quality commands the same price regardless of who is supplying the labor. Furthermore, each worker will be utilized so that his or her productive contribution is maximized, and the worker will be paid in accordance with his or her productive contribution. It is therefore a logical implication of the orthodox model that, except for a worker's unwillingness to work, his or her productive value will be fully utilized and fully rewarded.

The most obvious flaw in the neoclassical "proof" of equal opportunity is that the assumptions of the orthodox model are not valid. A second flaw is that the neoclassical definition of equal opportunity provides a limited concept of equal opportunity. It overlooks, for example, inheritance of wealth and the effect of socioeconomic origin on access to quality schooling.

The orthodox model may be amended so as to make it possible to deal with discrimination. Becker (1957), for example, conceived the notion that white workers demand higher pay if required to associate with black workers. Taking a different tack, but still maintaining the assumptions of perfect competition, Arrow (1973) and Stiglitz (1973) introduced into the neoclassical model the use of race on the part of hiring employers to determine worker quality. Although such amendments to the neoclassical model undermine the claim that competition ensures equal opportunity and equitable wages (Stiglitz, 1974), left intact is the focus on the characteristics of individual workers.[4] In addition, this form of analysis has the effect of redefining prejudice and bigotry as worker preferences and cost-efficient hiring criteria rather than phenomena of conflict among groups.

Ideological bias.– John Stuart Mill observed that the most unpleasant jobs were the worst paid. He criticized the compensatory principle for covering up this fact: "The inequalities of wages are generally in an opposite direction to the equitable principle of compensation erroneously represented by

4. Cf. the emphasis on lack of commitment to work on the part of women and blacks; for example, see Council of Economic Advisors (1973: chapter 4) and Feldstein (1973).

Adam Smith as the general law of the remuneration of labor" (Mill [1848], 1900:372).

Mill is saying two things. First, something remedial should be done about the low wages received by the poor. Second, orthodox economics inhibits the recognition of this particular social problem. Mill's concern for inequality and his criticism of orthodox economics for covering up inequality are typical of advocates of the segmentation approach (see Cain, 1976).

To Mincer (1970), Mill's observation is a stimulus to the cumulative theoretical development of orthodox economics. Mincer argues that the inequality in wages might be due to free choice among individuals: one person picks a career with constant wages over time, while another foregoes wages in the present in return for higher wages in the future.

> Some of the predications of the compensatory principle are often rejected *prima facie*: on the whole, occupations in which work is more unpleasant and unstable command lower, not higher wages. The costs of occupational training, however, can reconcile these apparent contradictions. (Mincer, 1970:3)

In contrast to Mill, Mincer appears to regard low wages as more of a theoretical problem than a social problem. It also appears that Mincer may be doing what Mill accused Adam Smith of doing, namely, developing ideas that inhibit the recognition of inequality as a social problem.

An important difference between the individualistic and segmental approaches is that the segmental approach has been favored by reformers and radicals, whereas the individualistic approach is favored by advocates of conservative policies. No doubt persons are drawn to one or the other approach on the basis of their ideological biases. More insidious is the tendency for the individualistic approach to lead to conservative social policies and the segmental approach to lead to interventionist social policies.

Policies suggested by the segmental approach include eliminating discriminatory hiring and improving the supply of jobs. From the individualistic perspective of orthodox economics, however, neither of these policies makes sense since competition among employers rules out discrimination, and job characteristics are not a subject of analysis. Furthermore, explaining inequality by the characteristics of individuals makes inequality seem natural and inevitable and suggests that change can only be accomplished by changing people. Moreover, the possibility of obtaining change through collective action is hard to conceive given the emphasis on competition among individuals.

Others have argued that the choice between theoretical approaches be made on the basis of predictive tests. (Friedman, 1953, holds that predictive tests are the sole criteria for judging a theory.) It is seldom possible in social science, however, simply to allow the facts to determine which approach is best (Roberts, 1974). Indeed, as Kuhn (1970) points out for the physical sciences, a theory can always be made more complicated so as to achieve consistency with the facts.

Rather than attempt predictive tests, this chapter compared approaches in labor economics and their effects on our awareness. The need for such a comparison stems from the intimate tie between social theories and social interests (see Mannheim, 1936; Mills, 1959).

The segmental approach directs attention beyond the individual in the search for the causes of and remedies for inequality. Since a labor segment is a set of jobs, labor segments focus attention on the characteristics of jobs as opposed to the characteristics of people. In contrast, individualistic analysis focuses on the characteristics of people. Thus it is easy to conceive from a segmentation perspective, but not from an individualistic perspective, that one of two otherwise identical persons could have a better job than the other.

A labor segment is also a group of people. Thus labor segments focus attention on the capacity of individuals to cooperate with one another in pursuit of collective goals. Hence it is easy to conceive from a segmentation perspective, but not from an individualistic perspective, that inequality among persons reflects the distribution of power among groups and classes rather than simply differences in the characteristics of individuals.

Labor segments provide a basis for defining intergenerational social mobility that avoids the confining tendencies of individualistic analysis. In addition, intergenerational mobility is a source of information on differences and relationships among segments and can help to define labor segments that reflect social structure.

CHAPTER THREE

CONTINGENCY TABLE ANALYSIS:

Effect of Origins on Destination Probabilities

SEGMENTATION models divide workers into labor segments. Once the labor segments are defined, the most straightforward approach to examining intergenerational male mobility is to construct a mobility table by cross-classifying the labor segments of father and son. This chapter examines such a mobility table constructed from the OCG data. The labor segments are the widely used twelve major census occupations, which were originally conceived of as socioeconomic categories (Edwards, 1934).

The correlation approach characterizes mobility with a single number. In contrast, the mobility table we will examine contains 144 cells. It seems clear, therefore, that the mobility table provides more information. Nevertheless, unless the information in the mobility table can be reduced to a smaller set of numbers, the correlation approach is likely to dominate due to its parsimony.

The percentage of sons in the same occupations as their fathers is a commonly used measure of immobility (for example, see U. S. Bureau of Census, 1964). This measure embodies a very narrow concept of what constitutes evidence that origins affect destinations. To emphasize the narrowness of this measure, the percentage of sons in the same occupation as their fathers shall be referred to as "strict occupational inheritance."

Rather than focus on the proportion of *all* sons who are in the same occupation as their fathers, we may examine strict occupational inheritance for each occupation. For which occupations are the chances of entry enhanced for persons whose fathers are in those occupations? Using the major census occupational categories, the first section shows that, with one exception, for all occupations there is a tendency toward strict occupational inheritance. On the average, the observed probability of entering a given occupation is roughly three times better for persons whose fathers are in that occupation.

The rate of strict occupational inheritance equals the proportion on the diagonal of the mobility table and thus is not affected by the pattern of the off-diagonal elements. The off-diagonal elements, however, provide further evidence of the effect of origins. A formerly common assessment of mobility tables, reached in several community studies, is that, although most sons are not in their father's category, long-distance mobility is rare: "Most mobility in both past and present has been mobility of relatively small degree" (Barber, 1957). In contrast, Goodman (1965) has argued for "quasi-perfect mobility," which holds that except for strict occupational inheritance, there is perfect mobility and hence only the diagonal of the mobility table is of interest.

The second section of this chapter examines the off-diagonal elements in order to address two fundamental questions: How equal are the probabilities of reaching the top of the socioeconomic structure? How equally shared is the risk of a destination at the bottom? To deal with these two questions attention is focused on the two columns of the mobility table depicting entry into the top and bottom occupations. This section provides evidence of "inheritance of poverty." For obtaining elite status, origins are shown to be even more important.

The data source is the mobility table published in the appendix of Blau and Duncan (1967:496). Unfortunately, there are data only for males. The data, which exclude men aged 20 to 25, are weighted for differential response rates so as to be more representative of the total population of 45 million males. The published table in Blau and Duncan is based on seventeen categories, which have been collapsed for the purpose of this chapter to the twelve major census occupations. The new table, in the form of transition probabilities, is appendix table A.1.

Strict occupational inheritance

The influence of origins is most apparent when father and son have the same occupation. Ratios of probabilities may be used to indicate the magnitude of this effect.

The second column of table 3.1 gives for each occupation the probability of strict occupational inheritance, the proportion of sons who are in their father's occupation:

$$P(j) = n(j,j)/n(j,\cdot) \quad (j = 1, \ldots, 12)$$

where $n(j,j)$ = number of observations with origin j and destination j.[1] The probabilities differ markedly, ranging from 10 percent for clerical workers and farm laborers, to 33 percent for salaried professionals. For the total sample, 21 percent of sons have the same occupations as their fathers; that is, 21 percent of the observations fall on the main diagonal of the mobility table.

The probabilities of strict occupational inheritance are not comparable since the occupations differ in size. For example, since only a very small proportion of workers are farm laborers (1.84 percent), it is not surprising that the probability of strict occupational inheritance is low for farm laborers.

To adjust for the effect of size on the probability of entering an occupation it is helpful to use as a comparison group persons who do *not* have fathers in the given occupation. The ratio of the probability of strict occupational inheritance to the probability for the comparison group shall be called the "odds favoring strict occupational inheritance," or, for brevity, "odds favoring inheritance." The odds favoring inheritance are directly interpretable as the number of times more likely it is for a person from the favored group to have a particular destination than it is for a person not in the favored group to have the same destination. Letting S(j) denote the probability of destination j for those who do not have origin j, the odds favoring strict occupational inheritance O(j) are defined as

$$O(j) = P(j)/S(j),$$

where

$$S(j) = [n(\cdot,j) - n(j,j)] \;/\; [N - n(j,\cdot)].$$

Table 3.1 displays the odds favoring inheritance alongside the P and S probabilities. As an example, since 10.2 percent of farm laborers' sons are farm laborers and 1.6 percent of those with other origins are farm laborers, the odds favoring inheritance are .102/.016 = 6.42. For sons of farm laborers, the probability of being a farm laborer is 6.4 times the probability for those with other origins.

1. A dot denotes summation, or marginal frequency:

$$n(j,\cdot) = \sum_{k=1}^{12} n(j,k).$$

TABLE 3.1. Strict occupational inheritance

	O(j)[a] *Odds favoring inheritance* (1)	P(j)[b] *Probability of strict inheritance* (2)	S(j)[c] *Probability of destination for persons of other origins* (3)
Nonmanual occupations			
1. Self-employed professionals	13.0	.175	.013
2. Salaried professionals	3.2	.334	.105
3. Salaried managers	2.6	.209	.082
4. Self-employed managers	2.5	.169	.068
5. Sales workers	3.2	.150	.048
6. Clerical workers	1.5	.096	.065
Manual occupations			
7. Craft workers	1.6	.294	.188
8. Operatives	1.5	.259	.173
9. Service workers	2.1	.111	.052
10. Laborers	2.4	.142	.060
11. Farmers	14.5	.178	.012
12. Farm laborers	6.4	.102	.016

a. $O(j) = P(j)/S(j)$

b. $P(j) = n(j,j)/n(j,\cdot)$

c. $S(j) = [n(\cdot,j) - n(j,j)]/[N - n(j,\cdot)]$

Despite their simplicity and ease of interpretation, the odds favoring inheritance, to my knowledge, have not been previously employed. The reason may be failure to separate strict occupational inheritance sharply from other aspects of the effect of origins. To adjust for the differential sizes of the categories of origin and destination, others have relied on the "indexes of association" first introduced by Rogoff (1953) and, independently, by Glass (1954). To obtain the indexes of association, the ratio of the observed frequency to the chance frequency is calculated for each cell of the mobility table. The chance frequency is the frequency predicted by the marginals under the assumption that there is no relationship between origins and destinations.

Elements on the main diagonal of the indexes of association matrix may be regarded as measures of strict occupational inheritance. (This matrix may be found in appendix table B.5.) The value corresponding to category j may be interpreted, like the odds of strict occupational inheritance, as the ratio of the probabilities of destination j for two different groups. The probability of strict inheritance is the numerator for both ratios, but the denominators differ. The denominator for the index of association is the probability of destination j averaged over all origins. (The average probability of destination j is the marginal probability of destination j, $n(\cdot,j)/N$.) Thus, the index of association provides a more ambiguous comparison since the two probabilities, unlike the odds favoring inheritance, do not refer to two complementary subsets of the observations.

In terms of P and S, the index of association for the j^{th} diagonal element is

$$R^*(j) = \frac{P(j)}{\lambda P(j) + (1-\lambda)S(j)}$$

where

$$\lambda = n(j,\cdot)/N.$$

Since P(j) is greater than S(j), the odds favoring inheritance are larger in magnitude than the indexes of association, giving a stronger impression of strict inheritance. A persuasive reason for not using the index of association to measure strict inheritance is that the value of this index depends on the size of the origin category: large λ, the proportion of the sample with origin j, will depress the index for given values of P(j) and R(j). Most dramatically, since almost 30 percent have farm origins, the index of association for the diagonal element referring to farmers is only 3.0, whereas the odds favoring inheritance for farmers is 14.5.

The odds favoring inheritance may also be compared to a measure proposed by Goodman (1969:5) which can be specialized to refer to strict inheritance. The "relative risk" of destination j (or "odds ratio") for those of origin j in comparison to those of other origins is equal to

$$\frac{P(j)}{1-P(j)} \div \frac{S(j)}{1-S(j)}.$$

The relative risk would give a more striking impression of strict inheritance, but its meaning may be less clear than a ratio of two probabilities. For example, the relative risk for persons of farm origins becoming farmers is 18.1, while the odds favoring inheritance are 14.5. (Goodman [1964] has examined alternative methods of estimating confidence intervals for the relative risk; these methods can also be used to assign confidence intervals to the odds favoring inheritance.)

The degree of strict inheritance is not the same for all occupations. As is evident from the first two columns of table 3.1, there is considerable variation in both the probabilities of strict inheritance and the odds favoring inheritance. The odds favoring inheritance, however, give a much stronger indication of the variability in the degree of strict occupational inheritance. The odds favoring inheritance range from a low of 1.5 for operatives and clerks, to 13.0 for self-employed professionals and a high of 14.5 for farmers. Origins are important in the second-order sense of determining the probability of strict occupational inheritance.

The simple average over the twelve occupations of the odds favoring inheritance is 4.5. Weighting by the proportions of sons in each occupation, the weighted average is 2.7. Weighting by the distribution of fathers yields a value of 5.8. As a conservative summary, it may be concluded that, on the average, having a father in an occupation multiplies the probability of entry into that occupation by a factor of three.

The odds favoring strict inheritance are especially high for the top of the occupational structure. Since access to the top has particular significance, the high inheritance at the top strongly contradicts the proposition that origins are not important.[2]

Entry into the top or bottom

Whether a son has the same socioeconomic position as his father is especially significant if the father is a member of the elite or in poverty. A position at the top of the socioeconomic structure is perhaps best indicated by exceptional wealth, power, or both; a position at the bottom by poverty. Since this study is based on broad occupational categories, this issue is approached by considering destination probabilities for the broad occupational categories of highest and lowest socioeconomic standing.

2. More specialized data confirm high odds favoring strict occupational inheritance at the top of the occupational structure. For example, using a 1961 national sample of male college seniors, Zelan (1964, 1972) determined that 35 percent of lawyers' sons planned a career in law whereas 5 percent of other seniors planned such a career. If we assume that lawyers' sons are twice as likely as all other persons to reach the senior year of college, these values suggest that for lawyers the odds favoring strict occupational inheritance are fourteen to one.

The category called self-employed professionals stands highest among the twelve census occupational categories on a variety of socioeconomic measures. For example, the median income for self-employed professionals, as of 1962, is 66 percent greater than the median income for salaried managers, which is the occupational category with the second highest median income (Blau and Duncan, 1967:27). Laborers, with a 1961 median income of \$2,189, make up the nonfarm occupation with the lowest socioeconomic standing.

From the odds favoring inheritance shown in table 3.1, it is evident that the chances of entering either of these occupational categories are enhanced by having a father in that occupational category. The odds favoring inheritance are 13.0 for self-employed professionals, by far the highest value for any of the ten nonfarm occupations. However, the odds favoring occupational inheritance do not indicate that sons of laborers are exceptional in the degree of strict occupational inheritance. The odds favoring inheritance, 2.4, are somewhat less than average for this occupation.

The odds favoring inheritance summarize the openness of an occupation by means of a single summary statistic. That statistic compares two probabilities of entry into an occupation. The probability of entry for persons with fathers in the occupation ("sons") is compared to the probability of entry for persons who do not have fathers in the occupation ("others"). Since "others" consists of eleven different origins, it is possible to compare separately the probability of entry for "sons" to the probability of entry for persons with each of the other eleven origins. The vector of odds against a particular occupational destination gives the probability of entry into that occupation for "sons" relative to the probability for persons with each of the other occupational origins.

The vector of odds against destination j consists of the elements

$$R(i,j) \qquad (i = 1, \ldots, 12),$$

where

$$\begin{aligned} R(i,j) &= T(j,j)/T(i,j) \\ &= P(j)/T(i,j), \end{aligned}$$

and

$$T(i,j) = n(i,j)/n(i,\cdot).$$

$T(i,j)$ is the probability of entry into occupation j for persons with origin i.

In comparison to the odds favoring inheritance, the vectors of odds against destinations provide a much less parsimonious summary of a mobility table. The vector of odds against a destination nevertheless does represent a summary of the information in a particular column of a matrix of transition probabilities. For each column of a 12-by-12 mobility table, there are 66 distinct pair-wise comparisons of the probabilities of entering a particular occupation (12 X 11/2 = 66). For example, we might compare the probability that an operative's son is a laborer with the probability that a salesman's son is a laborer (table A.1: row 8, column 10 *vs.* row 5, column 10). The vector of odds against a destination focuses on 11 of these 66 comparisons (row 1, column 10 *vs.* row 10, column 10; row 2, column 10 *vs.* row 10, column 10; etc.).

The second column of values in table 3.2 presents the odds against the laborer destination and the first column presents the corresponding column from the matrix of transition probabilities, which is appendix table A.1. (The values in the second column are found by dividing the value in row 10 of the first column by each value in the first column.) Each column of table 3.3 is a vector of odds against an occupational destination.

TABLE 3.2. The laborer destination

i	*Category of origin*[c]	T(i,10)[a] *Probability of laborer destination (%)*	R(i,10)[b] *Odds against destination*
1.	Self-employed professionals	2.75	5.15
2.	Salaried professionals	1.64	8.65
3.	Salaried managers	2.05	6.89
4.	Self-employed managers	1.82	7.77
5.	Sales workers	2.05	6.90
6.	Clerical workers	3.04	4.66
7.	Craft workers	4.81	2.94
8.	Operatives	7.54	1.88
9.	Service workers	6.27	2.26
10.	Laborers	14.16	1.00
11.	Farmers	8.50	1.67
12.	Farm laborers	13.41	1.06

a. T(i,10) = n(i,10)/n(i,•)

b. R(i,10) = T(10,10)/T(i,10) = P(10)/T(i,10)

c. Category of origin titles are abbreviated; complete titles are given in appendix table A.1.

TABLE 3.3. Matrix of odds against destinations, R[a]

Category of origin[b]	*Category of destination*[b] 1	2	3	4	5	6	7	8	9	10	11	12
1	1.00	1.00	2.02	3.64	1.32	2.28	4.63	5.84	5.82	5.15	8.44	12.05
2	5.09	1.00	1.55	3.39	1.87	1.20	3.04	2.04	3.13	8.65	20.69	59.09
3	4.61	1.37	1.00	2.00	1.90	1.17	1.97	3.66	6.93	6.89	33.48	66.93
4	4.45	2.36	1.08	1.00	1.58	1.48	2.19	2.81	3.72	7.77	14.98	24.88
5	6.47	1.99	1.10	1.55	1.00	1.57	2.46	2.49	3.43	6.90	10.59	69.53
6	7.21	1.31	1.71	3.05	1.92	1.00	1.73	2.81	1.82	4.66	12.83	*c*
7	17.76	2.79	2.32	2.27	3.17	1.24	1.00	1.48	2.15	2.94	22.38	27.79
8	20.00	3.10	3.72	2.57	3.42	1.45	1.23	1.00	1.87	1.88	19.08	10.90
9	21.95	3.60	2.66	2.67	2.63	1.02	1.40	1.24	1.00	2.26	17.06	41.41
10	64.65	5.96	4.82	4.62	4.15	1.20	1.30	0.98	1.22	1.00	14.60	9.01
11	26.07	7.24	4.73	2.38	6.08	2.05	1.49	1.27	2.11	1.67	1.00	2.39
12	84.40	16.09	6.71	3.88	7.60	2.51	1.43	1.00	1.37	1.06	2.86	1.00

a. R(i,j) = T(j,j)/T(i,j); T(i,j) = n(i,j)/n(i,•).

b. See table 3.2 for categories of origin and destination corresponding to numbers 1, 2, . . . , 12.

c. The numerator of the odds ratio is approximately zero.

The odds against the laborer destination are close to unity for persons with farm origins–1.06 for farm laborers and 1.67 for farmers. Since about 30 percent of the population have farm origins (matrix of inflow percentages, appendix table A.2, sum of elements 11 and 12 of marginal column), the rather typical value for strict occupational inheritance (2.4) is largely the result of the high rate of entry for persons with farm origins. In fact, about 42 percent of the laborers have farm origins (appendix table A.2, sum of elements 11 and 12 of column 10).

The impression that origins are strongly related to destinations emerges quite clearly from comparing probabilities of obtaining particularly favored or disfavored destinations. This impression, however, does not emerge if the observed destination distributions are examined in isolation. For example, although chances of becoming a self-employed professional are much higher for persons whose fathers are self-employed professionals, most sons of self-employed professionals are in occupations that have socioeconomic standing lower than their fathers' occupations.

The vectors for the self-employed professional and laborer occupations indicate that occupational origins have a major though varied impact on chances of ending up at either end of the occupational structure. Chances of being at the top show the greatest sensitivity to origins. For both occupations, there is a clear tendency for access chances to be lower for occupations more dissimilar in socioeconomic standing.

The generalization that long-distance mobility is less frequent than short-distance mobility does not fully describe the pattern of entry into either occupation. The probability of being a self-employed professional is very similar for sons of salaried professionals, salaried managers, or self-employed managers. In fact, there is a slight tendency among this group of occupations for the chances of entry to the top occupation to be higher for the lower standing occupations. The other two nonmanual occupations provide fewer chances for sons to rise to the top occupation but, compared with the drop-off in chances for craft workers' sons, there is striking uniformity of access to the top occupation for sons of nonmanual workers who are not self-employed professionals. The odds against the self-employed professional's destination range between 5 and 7 for the various nonmanual origins. There is also, except for nonfarm and farm laborers, fairly uniform access to the self-employed professional occupation for persons whose fathers were manual workers. For these five occupations the odds against the top destination range from 18 to 26. The odds against the top destination for very long-distance upward mobility are dramatic–65 for nonfarm laborers and 84 for farm laborers.

In sum, analysis of the probabilities of becoming a self-employed professional indicates (1) a strong advantage for "sons," (2) a barrier between the manual and nonmanual occupations to upward mobility, (3) an especially low chance for laborer's (farm and nonfarm) sons. Thus in describing access to the top there are four important categories of occupational origins: (1) self-employed professionals, (2) other white-collar workers, (3) laborers, (4) other manual workers. There is evidence of a manual-nonmanual barrier to downward mobility into the laborer occupation. The odds against entry are below 3 for all manual origins and above 4.5 for all nonmanual origins. Among the nonmanual origins there is little pattern to the risk of becoming a laborer; predictably, sons of clerical workers have a relatively high risk, but their chances of being a laborer are closer to those of the sons of self-employed professionals than to those of sons of craft workers. Among the manual occupations, the risk of becoming a laborer varies for persons who did not have a laborer father from roughly 1 to 3, with distance from the laborer occupation able to account for only the relatively low probability for craft workers' sons. The similarity of nonfarm laborers' and farm laborers' sons in their difficulty in entering the top occupation is reflected in their nearly identical risk of entering the laborer occupation. Surprisingly, service workers' sons have a lower risk of ending up in the bottom occupation than do sons of operatives or farmers. For purposes of describing the unevenness of the risk of ending up in the bottom occupation, the three groupings which appear to be of most use are these: (1) nonmanual, (2) laborers (including farm laborers), (3) other manual workers. Except for the absence of the self-employed professionals as a separate group, the aggregative groupings are the same ones mentioned above for summarizing the pattern of access to the self-employed professional occupation.

The risk of being at the bottom appears to be much more equally shared than the chances of being at the top, but occupational origins remain important. In particular, although occupational inheritance is not exceptionally high for laborers, the number of persons having farm origins who become laborers is so high that the risk of ending up a laborer is uniformly low for persons with nonmanual origins. Sons of laborers and farm laborers have virtually the same probability of ending up as laborers, which is about twice the probability for other manual origins, and about seven times the probability for nonmanual origins.

CHAPTER FOUR

CLUSTER ANALYSIS:

Divisions in the Socioeconomic Structure

SEGMENTATION models divide workers into groups or labor segments. Once the labor segments are defined, their relative socioeconomic standing emerges as an important issue. Data on intergenerational mobility may be used to determine relative standings of labor segments.

There are a number of ways to use mobility data to measure the socioeconomic distances from one labor segment to another. The pattern of distances between segments may be analyzed for evidence that segments form clusters which reflect underlying divisions in socioeconomic structure.

This chapter presents cluster analysis in theory, method, and data analysis. It discusses views of the social class structure of the labor force (theory). It examines the use of mobility data to establish the social distance between groups and then proposes a particular method of cluster analysis for aggregating occupations (method). Appendix C clarifies this method with a simple example based on a five-category social mobility matrix. It uses cluster analysis to determine the major socioeconomic groupings of workers (data analysis). As in chapter 3, the OCG mobility table for the major census occupational categories provides the data.

Theory

Behind the correlation model of intergenerational mobility lies an individualistic conception of socioeconomic structure. Each person is positioned at some point on a continuum. A person's position is an overall composite indication of productivity and/or prestige. The significance of the socioeconomic level of the father for the life-chances of the son is quantified by measuring the correlation of positions on the continuum of father and son. To interpret whether the correlation value is large or small, it may be squared and thereby transformed into a statement about the comparative size of two variances—the variance in position for all sons is compared with the predicted variance for all sons whose fathers had the same position on the socioeconomic continuum.

Chapter 3 emphasized the weaknesses of using the correlation to assess the strength of the relationship between the positions of father and son. Examination of the relative odds of entry into particular occupations, especially at the top and the bottom occupations, yields a strong impression that origins are an important determinant of life-chances. In chapter 5, the regression approach is further questioned because of the presence of nonlinearity and heteroscedasticity. These criticisms are easier to establish if the researcher thinks in terms of categories or groups rather than in terms of individuals. This chapter takes the focus on groups one step further by using the mobility data to define groups.

Following the example of Ossowski (1963), one can distinguish the problem of describing the class structure—the task of identifying the component socioeconomic groups of a society—from the debate over the nature of social classes. Regardless of the nature of social classes, mobility data may be examined for the presence of class cleavages. At the minimum, the mobility data may be used to establish nominally defined aggregative categories for subsequent use in summarizing the mobility data.

Previous research has examined the structure of the male labor force on the basis of mobility data. Among the earlier researchers there is some consensus that a mobility barrier separates an upper category from the rest (Landecker, 1960; Laumann, 1966; White, 1963). These researchers also are in some agreement concerning a secondary boundary separating a lower category of workers from the rest of the labor force (White, 1963; Laumann, 1966). More recently, Hope (1972), combining canonical correlation and cluster analysis, examined the Blau and Duncan data using their seventeen-category mobility table. He concludes that "the overall picture is one of a three-class society topped by an elite," with the divisions above a bottom farm category, between manual and nonmanual workers, and below the self-employed professionals. Blau and Duncan (1967) observe a split in the pattern of the ratios of observed to chance mobility between manual and nonmanual workers. However, they appear to claim that this represents a one-way barrier, which in no way inhibits upward mobility. Blau and Duncan, like Hope, do not find evidence for a boundary above a lower category of nonfarm manual workers; and, unlike Hope, Blau and Duncan do not observe a barrier isolating an upper category of workers.

In summary, there are four basic models of class structure which may be reflected in highly aggregative occupational clusters formed on the basis of mobility data: (1) isolated top, (2) isolated bottom, (3) undifferentiated middle, and (4) manual-nonmanual split. It should be realized that the data under consideration make it difficult to claim that tests of model (1) are related to the existence of an elite capitalist class. Model (2) seems to capture fairly well the liberal notion that the prime inequity in America is the existence of an underclass of poor, primarily black victims of the historical pattern of racial segregation (for example, Matza, 1966; Rainwater, 1968).

The choice between models (3) and (4) reflects the debate over the embourgeoisement thesis—that industrialism has led to the creation of a vast, homogeneous middle class. (For argument in favor of the embourgeoisement thesis, see Kerr et al., 1960; for an opposing view see Goldthorpe, 1964.)

The distinction between a middle class and a working class has been popular among sociologists and is typically defined on the basis of the manual-nonmanual distinction. For example, on the basis of "subjective" criteria such as values, attitudes, prestige, and self-identification, Kohn (1969) divides the major census categories into the working class and the middle class. Lipset and Bendix (1959) discuss objective as well as subjective evidence for basing the working class—middle class distinction on the manual-nonmanual distinction. Centers (1961) finds the working class—middle class split to be a part of most Americans' conception of social structure. Model (4) embodies the distinction between a middle and a working class.

Method

The social distance matrix.—The primary methodological task is to use the mobility table to derive a measure of the social distance separating each pair of occupations. Four methods are used in the data analysis.

The social distance between two occupations is reflected in the amount of mobility between the two occupations. Assuming a ranking of the occupations, the mobility can be divided into two types—upward mobility and downward mobility. Both amounts can be measured by means of a transition probability giving the percentage of those with a given origin occupation who have a given destination.

Since the amount of mobility to a destination reflects the percentage size of the destination, it is desirable to adjust for the destination sizes. The index of association achieves this adjustment for it is equivalent to dividing the transition probabilities by the destination probabilities. Rather than measure the closeness of two occupations, the reciprocal of the index of association may be used to measure social distance between two occupations.

The social distance between two occupations may be regarded as small if the two occupations are similar in their recruitment rates from the various origins. The social distance between occupations *i* and *j* is determined by the dissimilarity of columns *i* and *j* of an appropriate form of the mobility table. The matrix of inflow percentages is convenient since each column represents a probability distribution: the elements of column *i* give the percentage of the members of occupation *i* originating from each origin; column *i* gives the social composition of occupation *i* in terms of occupational origins. A simple approach to evaluate the dissimilarity of the two columns is to sum the absolute values of the differences in the elements, which are the differences in the rates of inflow from each origin:

$$d_{ij} = \sum_k |Y_{ki} - Y_{kj}|,$$

where

$$Y_{ij} = \text{Probability (origin=i|destination=j).}^1$$

Rather than compare columns, the distance between two occupations may be determined by comparing rows of the mobility table. Such a comparison of two rows indicates the similarity in destinations or occupational prospects for sons from the two different occupational origins. If the matrix of transition probabilities is used, then each row denotes the probability distribution of destinations for those of a particular origin. Thus, another approach to using mobility data to define social distance is to use the metric

$$d_{ij} = \sum_k |T_{ki} - T_{kj}|$$

where

$$T_{ij} = \text{Probability (destination=j|origin=i).}$$

To clarify the alternatives, let us consider the social distance between self-employed professionals and nonfarm laborers. The social distance ma-

1. As an alternative, the distance between two probability distributions may be found using an information measure (Kullback, 1968: especially 6–7, 110):

$$d_{ij} = \sum_k (Y_{ki} - Y_{kj}) \log \frac{Y_{ki}}{Y_{kj}}.$$

trix based on rates of upward mobility (appendix table B.3) places these two occupations at a distance of 5.8. This value indicates that the number of laborers' sons expected to be self-employed professionals is 5.8 times the observed number. (The expected number of laborers' sons is determined under the assumption of no association of origins and destinations.) The social distance matrix derived from the rates of downward mobility (appendix table B.4) indicates that the social distance between the two occupations is 2.4. Apparently, there is more movement downward between the two occupations than upward. The social distance between the two occupations on the basis of dissimilarity of origins (appendix table B.2) is 1.13 or 113 percent. This value indicates that in order for the social composition of the self-employed professionals occupation to be equalized to the social composition of the laborer occupation, 56.5 percent fewer self-employed professionals would be recruited from nonmanual occupations and 56.5 percent more from manual occupations. Based on dissimilarity of destinations (appendix table B.1), the social distance between self-employed professionals and laborers is 122 percent.

Assuming twelve occupational categories, the mobility table is a 12-by-12 asymmetric matrix containing 144 distinct elements. The set of social distances constitutes an 11-by-11 triangular matrix consisting of 66 distinct elements. The decision as to how to measure the social distance between occupations therefore yields a substantial reduction of the data. An additional step of data reduction remains to be taken.

The clustering procedure.—In contrast to the mobility matrix, a matrix of pairwise comparisons is particularly amenable to data reduction by application of geometrically inspired methods. Such methods of analysis are especially appropriate when the pairwise comparisons can be thought of as distances between points. (In the present case, the pairwise comparisons may be conceived of as mobility distances or social distances between occupations.) The interrelationship of a set of points may be revealed by dividing the points into clusters. The objective is to find "compact" clusters separated by substantial "distance."

A cluster configuration is a particular assignment of points to clusters, or more formally, the partitioning of a set of points into subsets. A central problem is to select criteria for evaluating a cluster configuration. A configuration may be evaluated on the basis of the compactness of the clusters: the closer the members of a cluster are to one another, the more compact is the cluster. Alternative terms for compactness include cluster size, spread, and diameter. A second basis for evaluating a configuration is the amount of separation of the clusters from one another. Thus, the objective is to find compact clusters separated by substantial distance.

Cluster compactness and cluster separation are two ways to look at the extent to which the clusters of a configuration command attention as distinct, individual entities. The term "individuation" will be used to refer to the overall assessment of the "goodness" of a cluster configuration, suggesting that the issue is the usefulness of regarding each cluster as an aggregative individual of some sort. The significance of compactness and separation is clarified if they are measured relative to one another. Thus, the individuation index may be usefully defined as the ratio of cluster separation to cluster size.

There are two views of the purpose of cluster analysis. Each view leads to different measures of cluster compactness and separation.

According to the first view, cluster analysis is primarily a technique of data reduction—a way of simplifying the data so that they are in a more manageable form for communication and comparison. A good cluster configuration is one that simplifies without sacrificing too much information. (Fisher, 1958 and 1969, develops this point of view from a decision-theoretic framework.) To implement this view, the mean within-cluster interpoint distance is used as the measure of cluster size,[2] and the mean between-cluster interpoint distance is the measure of cluster separation. The most common technique of cluster analysis used by proponents of the "information-loss view" is the minimization of within-cluster sum of squared errors (Fisher, 1958 and 1969; Ward, 1963).

Cluster analysis, according to the second view, is a way to discover "natural" groupings. This view is frequently applied in biology and psychology (Jardine and Sibson, 1971). The clusters are regarded as reflecting an underlying phenomenon rather than the need for simplification. A good cluster configuration would reflect regions of sparse observations and discontinuity. The clusters might be validated in terms of correlation with other phenomena. For spatial visualization from this point of view, the notion of neighbor is useful. An individual would not be a "natural" member of a cluster if its nearest neighbor were in a different cluster, although this fact might not be crucial from the information-loss view. To implement the natural-group view, the largest within-cluster distance between nearest neighbors is used as the measure of cluster size. The smallest distance between neighbors in different clusters is used to measure cluster separation. These two measures, introduced by Johnson (1967), who called them "single-link" measures, may appropriately be labeled "gap" measures. The gap measure of cluster size may be thought of as the largest gap that can be found within any of the clusters, whereas the gap measure of cluster separation indicates the gap between the two closest cluster boundaries.

For most criteria, the only way to obtain the optimum configuration is to search over all possible configurations. This approach is seldom feasible because of the incredibly large number of ways a set of points may be clustered. Sterling's number of the second kind, which for moderately large n is approximately m^{n-1}, gives the number of ways n points may be divided into m clusters (Anderberg, 1973). There are about 4,000,000 distinct configurations of four clusters that can be formed when there are only twelve occupations to cluster ($m^{n-1} = 4^{11} = 4{,}194{,}304$). Since it is not feasible to search over all possible configurations, a second-best procedure must be used to search for a near-optimum configuration.

There are vastly fewer ways to cluster a set of points if it is required that the clusters honor a given rank ordering of the points. A rank ordering would be violated, for example, if the first and third points are in the same cluster while the second point is in a different cluster. For n points and m clusters, the question is how many ways are there to select $(m-1)$ divisions from a set of $(n-1)$ possibilities. The answer is

$$\binom{n-1}{m-1} = \frac{(n-1)!}{(m-1)!\,(n-m)!}$$

Thus, there are only 165 ways to group twelve occupations into four clusters that will honor an initial rank ordering. The procedure used in the data analysis section searches over all configurations honoring the rank ordering implied by the occupational numbering scheme introduced in chapter 3.

The data analysis relies on the gap and the mean-distance criteria of compactness and separation. A corresponding individuation index, defined as the ratio of separation to size, is defined for each of the two approaches. Formulas for the individuation measures are given in table 4.1.

Data analysis

Occupational prospects. – Based on the social distance matrix reflecting differences in occupational prospects (appendix table B.1), the greatest gap in the occupational structure is between self-employed professionals and others. As a result, the best two-cluster configuration according to the gap individuation index, as indicated by the first line of table 4.2, consists of distinguishing a singleton cluster for the self-employed professionals occupation. Whereas over all other occupations the average distance to the nearest neighbor is 0.24 with a maximum of 0.28, the nearest occupation to self-employed professionals stands at a distance of 0.41. Thus, the gap measure of cluster size is 0.28, and the gap measure of cluster separation is 0.41. The gap individuation index, the ratio of these two numbers, indicates that the gap between the two clusters is about 50 percent larger than the gap within any of the clusters.

Examination of the matrix of transition probabilities (appendix table A.1) reveals that the exceptional social distance from the self-employed professionals occupation to any other occupation is largely due to the high occupational inheritance for this occupation. Almost 20 percent of self-employed professionals' sons took up their fathers' occupation, while there is no other occupational origin for which as many as four percent became self-employed professionals.

Among occupations which are adjacent according to the numbering scheme introduced in chapter 3, the largest gap is 0.50. This gap separates the

TABLE 4.1. Definition of individuation measures

	Gap	*Mean-Distance*
Cluster size	$\max_{ij \in W^*} d_{ij}$	$\text{Mean}_{ij \in W} d_{ij}$
Cluster separation	$\min_{ij \in B} d_{ij}$	$\text{Mean}_{ij \in B} d_{ij}$
Individuation index	$\dfrac{\min_{ij \in B} d_{ij}}{\max_{ij \in W^*} d_{ij}}$	$\dfrac{\text{Mean}_{ij \in B} d_{ij}}{\text{Mean}_{ij \in W} d_{ij}}$

W is the set of within-cluster pairs. B is the set of between-cluster pairs. W and B are complementary subsets of the set of *ij* pairs where *i* and *j* denote points to be clustered.

W = ⟨ij; point *i* and point *j* in same cluster⟩.

B = ⟨ij; point *i* and point *j* in different clusters⟩.

W* is a subset of W and consists of all within-cluster "nearest neighbors." Point *j* is the nearest neighbor to point i if no other point is closer to point *i*.

W* = ⟨ij; $ij \in W$ and $d_{ij} < d_{il}$ and $d_{ij} < d_{kj}$, for any *k* and *l* belonging to the same cluster as *i* and *j*⟩, where

d_{ij} = distance between point *i* and point *j*.

TABLE 4.2. Cluster Analysis Based on Dissimilarity in Destinations: Description of Cluster Configurations with Best Individuation

Number of Clusters	Criterion Index	Cluster Configuration	*Index of Individuation* Ratio of Separation to Size		*Size of Clusters* Distance Within Clusters		*Cluster Separation* Distance Between Clusters	
			gap	*mean*	*gap (max)*	*mean*	*gap (min)*	*mean*
2	Gap	1,2–10	1.47	1.46	.28	.55	.41	.80
2	Mean	1–6, 7–10	1.21	1.98	.41	.39	.50	.78
2	Avg. value of indexes		.74	1.49	.42	.50	.29	.73
3	Gap	1, 2–6, 7–10	1.47	2.26	.28	.33	.41	.75
3	Mean	dd.	dd.		dd.		dd.	
3	Avg.		.58	1.58	.43	.45	.24	.69
4	Gap	1, 2, 3–6, 7–10	1.06	2.26	.26	.31	.28	.70
4	Mean	dd.	dd.		dd.		dd.	
4	Avg.		.53	1.66	.44	.41	.23	.66
5	Gap	(3456) (89 10)	.96	2.21	.26	.30	.25	.67
5	Mean	(345) (789 10)	.89	2.34	.30	.29	.26	.67
5	Avg.		.51	1.72	.44	.39	.22	.65

6	Gap	(45) (789 10)	1.05	2.31	.25	.28	.26	.66
6	Mean	(345) (789)	.82	2.49	.30	.26	.24	.65
6	Avg.		.51	1.77	.43	.37	.21	.63
7	Gap	(45) (89 10)	1.05	2.56	.24	.25	.25	.63
7	Mean	(45) (789)	.96	2.65	.25	.24	.24	.63
7	Avg.		.52	1.83	.42	.35	.21	.62
8	Gap	(45) (89)	1.04	2.87	.23	.21	.24	.62
8	Mean	dd.	dd.		dd.		dd.	
8	Avg.		.56	1.88	.39	.34	.20	.61
9	Gap	(45)	1.19	3.1	.20	.20	.23	.61
9	Mean	dd.	dd.		dd.		dd.	
9	Avg.		.67	1.99	.33	.33	.20	.60

Note: (345) (789) is the same as 1, 2, 3–5, 6, 7–9, 10. Nonfarm data.

bottom nonmanual and the top manual occupations. Thus, a manual-nonmanual division of the occupational structure yields the best configuration in terms of the gap measure of cluster separation. The gap measure of cluster size is very large because of the distance between the self-employed professionals occupation and its nearest neighbor, and therefore the gap individuation index has a low value. The manual-nonmanual division, however, minimizes information loss in terms of the average of the within-cluster distances and maximizes the mean-distance individuation index. For the configuration consisting of a manual and nonmanual cluster, the average distance between points in separate clusters is twice the average distance between points in the same cluster.

Considering which two of the two-cluster configurations have the best individuation, the best three-cluster configuration is not unexpected. The best three-cluster configuration consists of self-employed professionals, other nonmanual workers, and manual workers. The gap individuation index and the mean-distance individuation index are both maximized for this configuration over all three-cluster configurations. The gap index for this configuration, at 1.5, equals its two-cluster maximum. The mean-distance index increases in value from its two-cluster maximum of 1.5 to 2.3.

Distinguishing a fourth cluster does not produce a configuration which is as well individuated as the best three-cluster configuration. The value of the mean-distance index for the best four-cluster configuration equals the three-cluster maximum of 2.3, but the gap index falls about 50 percent from its previous maximum to 1.1. The best four-cluster configuration and the best configurations on either the mean-distance or gap index for all higher configuration levels are found by subdivision of the best three-cluster configuration.

Compared with the best three-cluster configuration, the best four-cluster configuration distinguishes salaried professionals as a singleton cluster. The exclusion of the salaried professionals from the nonmanual cluster reduces the between-cluster separation to the distance between salaried professionals and clerical workers. The fact that salaried professionals, in terms of occupational prospects of their sons, are similar to the lowest ranked of the nonmanual occupations suggests that the salaried professionals be kept in the nonmanual cluster.

There is some support for regarding a bottom category as the best candidate for the fourth cluster. The two greatest interpoint distances in the manual cluster reflect the inferior prospects for laborers' sons relative to prospects for service workers' sons and craft workers' sons. Examination of the transition probabilities indicates that the dissimilarity in occupational prospects is most apparent in rates of entry into the different nonmanual occupations. For example, laborer and crafts occupations clearly belong to the same cluster in terms of proportion of sons who enter the crafts occupation, but the proportion of sons who become salaried professionals is very different for the two occupations.

The gap individuation index continues to deteriorate with an increase in the number of clusters. Thus, the best configuration is at the three-cluster level. It is helpful, nevertheless, to start at the highest clustering level and work down to lower clustering levels to clarify the basis of the main divisions in the occupational structure.

The manual cluster begins with the aggregation of operatives and service workers. The nonmanual cluster starts by combining self-employed managers and sales workers. Added next to the nonmanual cluster are salaried managers, who are followed by clerical workers. It is ambiguous whether craftsmen or laborers join the manual cluster next. Despite the delay in the assignment of the occupation of clerical worker to the nonmanual cluster, and the occupation of craft worker to the manual cluster, there is no clustering level for which the best configuration includes a cluster in which both of these occupations are members. More generally, despite the heterogeneity of the manual and nonmanual clusters, there is little support for an undifferentiated middle cluster that includes both manual and nonmanual workers.

Four specifications of social distance

The cluster analysis was repeated for three other specifications of the social distance matrix. The four different measures of the social distance separating occupations *i* and *j* are:

(1) dissimilarity in destinations for those with origins *i* and *j* (dissimilarity in social prospects or life chances) (appendix table B.1);
(2) dissimilarity in origins for those with destinations *i* and *j* (dissimilarity in social origins or social composition) (appendix table B.2);
(3) upward mobility rate between occupations *i* and *j* (appendix table B.3);
(4) downward mobility rate between occupations *i* and *j* (appendix table B.4).

For each of the four specifications of social distance, the analysis was done with and without farm observations.

Table 4.3 presents the best configurations for the two-cluster, three-cluster, and four-cluster levels on the basis of the two individuation indexes. Focusing on configurations with a small number of clusters limits attention to the more salient divisions in the occupational structure. Table 4.3 contains forty-eight entries. There is one entry for each combination of choices arising from the four specifications of social data, the gap index *vs.* the mean-distance index, farm data included *vs.* nonfarm data, and the three clustering levels.

None of the forty-eight alternatives for best configuration contains a middle-range cluster encompassing both manual and nonmanual occupations. Instead, for most configurations, there is a division between manual and nonmanual occupations. This division is clearest on the basis of occupational prospects and downward mobility. On the basis of social composition, it is unclear whether to assign the clerical occupation to the manual or to the nonmanual cluster. And based on upward mobility the craft occupations sometimes appear to belong to the nonmanual cluster. The craft and clerical occupations tend to resist assignment to either cluster, and they tend not to be close to one another as well.

The second most consistent division in the occupational structure is the farm boundary. This feature is most clearly evidenced by differences in social composition: few persons with nonfarm origins end up on the farm. The pattern of upward mobility from farm origins, however, is not strikingly dissimilar to the rates of upward mobility for other manual workers, especially service workers and laborers.

The social isolation of the very top occupation is the third most salient feature. This feature is most evident on the basis of occupational prospects and least evident on the basis of direct mobility flows. There is a lack of support for a cluster made up of supervisory and professional workers.

A final feature of the occupational structure, but one which is not clearly distinguished by the methods of cluster analysis used, is the isolation of a bottom category of workers. An isolated (nonfarm) bottom category does not show up in any of the entries based on differences in occupational prospects and social composition when attention is restricted to below the five-cluster configuration level. An isolated bottom category does show up on the basis of upward or downward mobility flows. However, there is support for several different definitions of the bottom category: laborers only; laborers and service; laborers, service, and operatives.

The first of the two parts of the data analysis of this chapter discussed in detail the results of determining social distance on the basis of differences in occupational prospects, with attention limited to nonfarm data. The second

TABLE 4.3. Cluster Configurations with Best Individuation

	Dissimilarity in destinations		*Dissimilarity in origins*		*Upward mobility*		*Downward mobility*	
	Nonfarm	*All*	*Nonfarm*	*All*	*Nonfarm*	*All*	*Nonfarm*	*All*
2-Cluster Configurations								
Gap Index	1, 2–10	1–6, 7–12	1, 2–10	1–10, 11–12	1–7, 8–10	1–7, 8–12	1–6, 7–10	1–10, 11–12
Mean Index	1–6, 7–10	dd.	1–5, 6–10	1–10, 11–12	1–6, 7–10	1–6, 7–12	dd.	dd.
3-Cluster Configurations								
Gap Index	1, 2–6, 7–10	1, 2–6, 7–12	1, 2–5 6–10	1–10, 11, 12	1–6, 7, 8–10	1–6, 7, 8–12	1–6, 7, 8–10	1–7, 8–10, 11–12
Mean Index	dd.	dd.	1, 2–6, 7–10	1–5, 6–10, 11–12	dd.	dd.	dd.	1–6, 7–10, 11–12
4-Cluster Configurations								
Gap Index	1, 2, 3–6, 7–10	1, 2–10, 11, 12	1, 2, 3–5, 6–10	1, 2–10, 11, 12	1–6, 7, 8, 9–10	1–6, 7, 8, 9–12	1–6, 7, 8–9, 10	1–6, 7, 8–10, 11–12
Mean Index	dd.	1, 2–6, 7–10, 11–12	1, 2–5, 6, 7–10	1, 2–5, 6–10, 11–12	dd.	dd.	1–5, 6, 7, 8–10	dd.

Note: The columns refer to different specifications of the social distance matrix.

part examined the robustness of the results to different specifications of the social distances among occupations and to the inclusion of farm data. The basic conclusion of part one is that the ten nonfarm occupations divide into three main clusters: (1) a singleton cluster consisting only of self-employed professionals, (2) a cluster consisting of the five other nonmanual occupations, and (3) a cluster made up of the four manual occupations. There is also a lesser degree of evidence for distinguishing an isolated bottom category. Part two provides further support for these results but indicates an isolated farm category as well and raises some questions about the exact location of the cluster boundaries.

CHAPTER FIVE

REGRESSION ANALYSIS:

Accounting for Differences in SEI among Workers

AN IMPORTANT characteristic of any research into socioeconomic inequality is whether the individual or the group is the basic unit of analysis. That is, underlying any study of socioeconomic inequality is an implicit decision to conceptualize and measure the causes and extent of inequality either in terms of individuals or in terms of groups. As argued in chapter 2, there are strong advantages to focusing on groups for developing a realistic understanding of socioeconomic inequality that will be helpful in predicting and in devising distributional policies. Furthermore, rather than define groups simply on a demographic basis, it is useful to define socioeconomic groups that reflect the institutional segmentation of the labor market into separate occupations. The mobility table analysis in chapter 3 and the cluster analysis in chapter 4, motivated by this interest in socioeconomic groups, use data on intergenerational mobility to characterize occupations.

The individual is the basic unit of analysis in orthodox economic theory and American ideology. This individualistic perspective is applied by those researchers who have argued that modern methods of empirical analysis require us to conclude that socioeconomic origins are not very important in America. Rather than ignore this approach or merely to point out its limitations, there is value in translating our ideas into the individualistic framework in order to clarify our research for persons preferring to think in individualistic terms.

There are other reasons for individualistic analysis of mobility data. The data are collected as observations on individuals; even if focus is to be on groups, individual observations must be examined so that they can be appropriately classified. There are difficulties with categorical data that may be mitigated if continuous variables defined on individuals are used. Heterogeneity is likely to be an acute difficulty when a small number of categories is used. If a large number of categories is used, very large samples may be required, and the large number of categories may make the analysis intractable and difficult to summarize.

To talk in a statistical way about origins affecting destinations amounts to saying that the probability distribution of destinations depends on origins. The analysis in chapters 3 and 4 was largely based on pairwise comparison of such probability distributions. If destinations are determined by a continuous measure of socioeconomic position, then it is easier to characterize the effect of origins on destinations. The effect can be described by (1) the effect of origins on the expected value of the continuous measure of socioeconomic position, or the effect on the location of the destinations probability distribution; and (2) the difference between the dispersion in socioeconomic position for the entire population and the dispersion holding origins constant.

This chapter uses the individual as the basic unit of analysis. The analysis is oriented toward accounting for differences among individuals. The primary focus is on the Duncan socioeconomic status index (SEI), but some attention is also given to explaining individual differences in education and income. This focus on individual differences involves viewing persons as arrayed along a continuum rather than as members of groups and represents a departure from the point of view of the previous chapters in which occupations are the basic units of analysis.

The segmentation perspective, nevertheless, is evident in this chapter. Whereas destinations are handled using a continuum, origins are handled using broadly defined socioeconomic categories. Both a twelve-category and a four-category scheme are used. This asymmetrical treatment of origins and destinations also characterizes some of the qualitative empirical research on social stratification. For example, Hollingshead (1975) describes the social origins of adolescents in a particular community by means of a scheme of five distinct social classes, but the adolescents are treated as having infinite gradations of difference from one another according to a number of status dimensions.

The application of the segmentation perspective leads to the problem of heterogeneous categories. Origin categories are invariably heterogeneous in terms of the socioeconomic standing of fathers. Let us consider, as an example, the origin category defined by self-employed professionals. One father may be a wealthy corporate lawyer and another a poorly paid itinerant musician. On the basis of variance in father's years of schooling, there is in the data to be discussed considerably more heterogeneity among fathers who are self-employed professionals than there is among the entire population of fathers. One reason for differences in the socioeconomic standing of sons of self-employed professionals is that their fathers differ along socioeconomic dimensions. Information on the heterogeneity of fathers for each category of origin is needed in order to judge how similar the sons are in terms of origins even though they are classified as having the identical origin.

The importance of father's income and education may depend on the category of origin. That is, sensitivity of sons to within-category differences may depend on category of origin. For example, nearly all self-employed professionals are able to provide good schooling opportunities for their sons. This homogenization effect may be attributable to a similarity in work experience or to social class. It is only by distinguishing heterogeneity among fathers from heterogeneity among their sons that we can consider this cause of differential dispersion across origins in the distributions of destinations. Thus, it is particularly interesting, in light of the heterogeneity of their fathers, that we shall find that, relative to other origins, sons of self-employed professionals are unusually homogeneous. The variance in the education index for fathers who are self-employed professionals is over 3.5 times the variance in the education index for their sons. Thus, in years of schooling, sons of self-employed professionals are far more homogeneous than their fathers.

Effect of origins on expected SEI

An obvious starting point is to assume that son's expected SEI is a linear function of father's SEI:

$$E(Y|X) = \alpha + \beta X \qquad [5.1]$$

where X = father's SEI and Y = son's SEI.

The linearity assumption is implicit in the use of the father-son SEI correlation coefficient to measure (im)mobility. The linearity assumption is also implicit in the use of various statistics based on the fitted linear regression of son's SEI on father's SEI–including the regression slope coefficient, R^2, $\bar{R}^2$, the variance in son's SEI holding socioeconomic origins constant, and the standard error of estimate.

The least squares fit[1] for model 5.1 is:

$$Y = 26.4 + .486X.$$

The standard error for the slope is .015, $\bar{R}^2$ is 16.7 percent, and the standard error of estimate (SEE) is 22.5.

A considerably more general model of the effect of father's SEI on son's SEI is:

$$E(Y|X;k) = \alpha_k + \beta_k X \quad (k = 1, \ldots, K) \qquad [5.2']$$

where *k*, denoting category of origin, is determined by information on father's socioeconomic position. To facilitate analysis of the effect of the category of origin, the α_k parameters may be redefined as deviations and the model rewritten:

$$E(Y|X;k) = E(Y) + \alpha_k + \beta_k [X - E(X)]\,. \qquad [5.2]$$

If father's SEI equals the expected SEI for all fathers, then by definition the expected SEI for sons with origin *k* is α_k units above the expected SEI for all sons.

Model 5.1 treats socioeconomic position of both father and son as points on a continuum. Model 5.2 uses a continuum only for the son's position. In addition to rejecting a continuum specification of origins, model 5.2 incorporates a segmentation perspective by allowing for the parameters of the relationship between origin and destination to depend on the category of origin. This may be viewed as an instance of a general proposition emanating from the segmentation perspective: whatever differences there are among individuals in resources, the effect of these differences depends on the segment; or, more generally, the process whereby life chances are determined is different for different segments. (A still more general segmentation model would allow the parameters to depend not only on the category of origin but also on categories of destination.) Model 5.2 will be referred to as the "general segmentation model."

With model 5.1, the simple linear regression of son's SEI on father's SEI, as a standard of comparison, this section successively considers three versions of the general segmentation model. In the first version, all β_k are constrained to be zero; in the second version, the β_k are constrained to be equal; and in

1. The sample is all observations from the OCG dataset that meet these conditions: (a) age of son at time of interview is between 35 and 44, inclusive; (b) occupations of both father and son are specified. The sample size, *N*, equals 4,815.

the third version, no restrictions are placed on the β_k. The first version, called the polar segmentation model, assumes identical socioeconomic position for all fathers in the same socioeconomic category. In the second version, heterogeneity is acknowledged, but sensitivity of sons to the within-category socioeconomic differences among fathers is assumed to be constant. The general version allows the sensitivity to depend on the category of origin.

The origin categories are operationalized using father's major census occupation. The twelve major census occupations distinguish broadly defined types of work, distinguish manual from nonmanual employment, are easy to apply to the data, have been widely used, and are central to the arguments of chapters 3 and 4. The categories of origin are not defined by dividing the range of fathers' SEIs into intervals. If the categories of origin were so defined, then comparison of models 5.1 and 5.2 would be a matter of investigating the linearity of the regression of son's SEI on father's SEI.

Least-squares fit of the polar segmentation model.—We first examine the least-squares fit for the polar version of the general segmentation model:

$$E(Y|k) = E(Y) + \alpha_k \qquad [5.2a]$$

where Y = son's SEI and k = index for father's major census occupation.

The main features of this model are these: (a) destinations, or socioeconomic positions of sons, are measured by son's SEI; (b) segments of origin are defined by father's major census occupation; and (c) differences in socioeconomic origins among persons with the same category of origin do not affect expected SEI (all $\beta_k = 0$).

The polar segmentation model is fitted by simply finding the mean SEI for persons grouped by segment of origin and then subtracting each group mean from the grand mean. The deviations of the twelve SEI means for males grouped by father's major census occupation, labeled α_k, are displayed in table 5.1. (A major appeal of the polar segmentation model is that it is estimated in such a straightforward fashion. The group means are the least-squares estimates for the model. The group means, however, are intuitively meaningful in the absence of the theorems justifying least-squares estimation. In addition, unlike models involving continuous independent variables, robust methods, as we shall see, are very easy to apply to the polar segmentation model.)

The effect of segment of origin on mean SEI is clarified by expressing the magnitudes in percentile units. The second column of table 5.1 shows for each segment of origin the number of percentiles by which son's mean SEI exceeds the median for all persons. For example, the clerical worker's seg-

TABLE 5.1. Mean SEI of males grouped by origin

k	*Father's major census occupation*	*Mean minus grand mean* $(\hat{\alpha}_k)$	*Percentiles deviation from grand median*	*Standard error* $SE(\hat{\alpha}_k)$
1	Self-employed professionals	27.0	33.	3.6
2	Salaried professionals	21.5	27.	2.8
3	Salaried managers	17.6	21.	2.6
4	Self-employed managers	11.6	13.	2.4
5	Sales workers	15.7	22.	2.6
6	Clerical workers	16.2	22.	2.7
7	Craft workers	1.6	6.	2.2
8	Operatives	− 2.1	2.	2.2
9	Service workers	0.6	5.	2.6
10	Nonfarm laborers	−. 6.0	− 1.	2.4
11	Farmers	−10.7	− 6.	2.2
12	Farm laborers	−15.0	−10.	2.1

$\hat{E}(Y) = 39.35$; $N = 4{,}815$; SEE = 22.48; $\bar{R}^2 = 17.11\%$

$\hat{\alpha}_k = \bar{Y}_k - \bar{Y}$, $\bar{Y}_k = \sum_i Y_{ki}/N_k$, $Y = \sum_{ki}\sum Y_{ki}/N$, where Y_{ki} = SEI of *i*th person with origin *k*.

ment has a value of 22 and the craft worker's segment a value of 6; these values indicate that the mean SEI for sons of clerical workers is 22 percentage points above the median (the 72nd percentile), whereas the mean SEI for sons of craft workers is only 6 percentage points above the median (56th percentile).

The fitted polar segmentation model provides information on the relative socioeconomic standing of the major census occupations. The twelve SEI means provide support for a clustering of occupations into a nonmanual, a blue-collar, and a farm cluster. To a lesser degree, there is also evidence for the isolation of the top and bottom of the occupational structure: mean SEI is especially high for sons of self-employed professionals and low for sons of laborers.

The twelve SEI means predict individual SEI very slightly better than the linear regression of son's SEI on father's SEI. The value of $\bar{R}^2$ is raised from

16.7 to 17.1 percent. In demonstrating the association between origins and destinations, the SEI construct may be dispensed with for purposes of measuring origins. That is, a much simpler method of measurement may be used to demonstrate, or to control for, the effect of origins.

In using model 5.2a rather than model 5.1, one may be less likely to mistakenly assume that the full effect of origins has been captured. Whereas one might be lulled into assuming that socioeconomic origin is held constant by controlling for father's SEI, it is easy to point out, for example, that all persons whose fathers were self-employed managers do not have identical origins. By showing that the $\bar{R}^2$ from the regression of son's SEI can be duplicated by simply regressing son's SEI on a set of dummy variables for the major census occupations, we have clarified the need for an upward revision in the $\bar{R}^2$ statistic.

Not only does model 5.2a demonstrate that the $\bar{R}^2$ from the linear regression of son's SEI on father's SEI understates the strength of association between origin and destination; the coefficients for model 5.2a indicate that the linear regression obscures the clustering of occupations into nonmanual, blue-collar, and farm workers. Model 5.1, it will be recalled, does not raise the issue of relative socioeconomic standing of segments; the estimates of the polar segmentation model implied by model 5.1 also overlook, erroneously, the occupational clusters.

The parameters of the polar segmentation model may be estimated under the assumption of model 5.1 that there is a linear relationship between SEI of father and son. For each category of origin, mean son's SEI is replaced by the predicted mean obtained by substituting mean father's SEI into the fitted regression of son's SEI on father's SEI:

$$\hat{\bar{Y}}_k = a + b\bar{X}_k,$$

where

$\bar{X}_k$ = mean father's SEI for origin k,
$\hat{\bar{Y}}_k$ = predicted mean son's SEI,
a = fitted least-squares intecept, and
b = fitted least-squares slope.

Table 5.2 compares the calculated SEI means with the means predicted by the assumption of a linear relationship.

The occupational clusters are obscured by the linearity assumption since within the set of nonmanual occupations and within the set of blue-collar occupations, there is greater homogeneity in mean son's SEI than predicted

TABLE 5.2. Use of father's SEI to predict SEI means for sons grouped by origin

k	*Father's major census occupation*	*Mean minus grand mean* ($\hat{\alpha}_k$)	*Predicted mean minus grand mean* ($\hat{\alpha}_k^*$)	*Difference* ($\hat{\alpha}_k - \hat{\alpha}_k^*$)
1	Self-employed professionals	27.0	26.5	0.5
2	Salaried professionals	21.5	22.0	–0.5
3	Salaried managers	17.6	19.6	–2.0
4	Self-employed managers	11.6	9.7	1.9
5	Sales workers	15.7	11.6	4.1
6	Clerical workers	16.2	10.2	6.0
7	Craft workers	1.6	2.1	– .5
8	Operatives	– 2.1	– 5.4	3.3
9	Service workers	0.6	– 3.8	3.2
10	Nonfarm laborers	– 6.0	– 9.6	3.6
11	Farmers	–10.7	– 6.3	–4.4
12	Farm laborers	–15.0	– 9.0	–6.0

$\hat{\alpha}_k^* = \bar{Y}_k - \bar{Y}$. $\hat{\alpha}_k = \hat{Y}_k - \bar{Y}$, where $\hat{Y}_k = 26.36 + .481\bar{X}_k$.

by the linear regression. Underpredicted is the SEI gap between the bottom nonmanual and the top manual category. Also underpredicted is the differential in mean son's SEI between the farm and the nonfarm categories.

Other applications of the polar segmentation model. – The twelve SEI means provide information on the relative socioeconomic standing of occupations. The results suggest that simple linear regression of son's SEI on father's SEI understates the effect of origins. Reestimation of the polar segmentation model by a more robust method strengthens this evidence.

In place of means, SEI medians may be used to fit the polar segmentation model. Table 5.3 displays the estimates of the effect of origin category on expected SEI based on the medians, which are the minimum absolute error (MAE) estimates for the polar segmentation model. (Medians minimize the sum of absolute errors whereas means minimize the sum of squared errors.) The second column of values in the table is the differences between the

TABLE 5.3. Effect of father's major census occupation on median SEI

k	*Category of origin*	$\hat{\alpha}_k$	*Median-mean*
1	Self-employed professionals	35	8.0
2	Salaried professionals	28	6.5
3	Salaried managers	24	6.4
4	Self-employed managers	15	3.4
5	Sales workers	24	8.3
6	Clerical workers	24	7.8
7	Craft workers	2	0.4
8	Operatives	− 3	−0.9
9	Service workers	2	1.4
10	Nonfarm laborers	− 6	0.0
11	Farmers	−18	−7.3
12	Farm laborers	−14	1.0

estimates of the α_k based on medians and the estimates based on the means as given in table 5.1.

Compared to the means, the medians show a substantially greater influence for socioeconomic origins. The estimated differences across origins in the location of the SEI probability distributions are revised upward by several points for all six nonmanual categories. On the average, for the nonmanual categories of origin, son's expected SEI is increased by 6.7 points. The expected SEI for the other origins is about the same, except for a downward revision of 7.3 points for the farmer origin.

The least-squares estimates of model 5.1 and model 5.2a produced an $\bar{R}^2$ of about 17 percent. Under the assumption of normality, such a value of $\bar{R}^2$ would lead us to expect that the conditional mean absolute deviation holding origins constant is 8.9 percent less than the unconditional mean absolute deviation.

The unconditional mean absolute deviation (or mean absolute error) is the mean absolute deviation around the overall median son's SEI, and the conditional mean absolute deviation is the mean absolute deviation around the twelve conditional medians. The conditional medians equal the α_k given in table 5.2 increased by the overall median, which is equal to 37. The value of the unconditional mean absolute deviation is 21.3, and the conditional mean absolute deviation is 18.3. Thus, the calculated proportional reduction in the mean absolute deviation for the estimates based on the medians is 14.1 percent, which is 60 percent higher than anticipated.

The descriptive value of the polar segmentation model is confirmed for other continuous measures of socioeconomic destination. Two alternatives which are available in the OCG data are income and years of schooling, both of which are indexes. Each index ranges over the integers from 0 to 8, with each value representing an interval.[2] By linear interpolation, the estimates were converted into dollars and years of schooling.

Father's major census occupation consistently outperforms father's SEI. Using mean son's SEI for the twelve categories of origin to predict son's SEI fits slightly better than the regression of son's SEI on father's SEI, with $\bar{R}^2$ rising from 16.7 to 17.1. For predicting income, the linear regression of son's income on father's SEI produces an $\bar{R}^2$ of 10.6, while the income means yield an $\bar{R}^2$ of 12.9. Using the son's years-of-schooling means rather than the linear regression of son's years of schooling on father's SEI raises $\bar{R}^2$ from 18.4 to 20.2. (The income calculations are for the subsample of 3,338 for which son's income was reported and was not 0. For this subsample, using means rather than linear regression increases the $\bar{R}^2$ for SEI from 16.1 to 17.2 and increases the $\bar{R}^2$ for education from 18.5 to 20.2.)

The manual-nonmanual distinction is most salient in the means of sons' SEIs, also clear in the education means, but far less evident in income means. To clarify the effect of the choice of socioeconomic destination measure it is helpful to examine the difference between the means of the lower nonman-

2. Following are the definitions of each index:

Value	*Income index (in dollars)*	*Education index (years)*
0	0	None
1	1– 1,999	Elementary, 1–4
2	2,000– 2,999	Elementary, 5–7
3	3,000– 3,999	Elementary, 8
4	4,000– 4,999	High School, 1–3
5	5,000– 6,999	High School, 4
6	7,000– 9,999	College, 1–3
7	10,000–14,999	College, 4
8	15,000+	College, 5+

ual origins and the top manual origin as a percentage of the range for the ten nonfarm means: a large percentage difference is indicative of at least the descriptive value of the manual-nonmanual distinction. The manual occupation of highest socioeconomic standing is unambiguously that of craft workers. It is ambiguous whether the lowest standing nonmanual occupation is self-employed managers, sales workers, or clerical workers.

Comparing mean SEI for sons in clerical occupations with the mean for sons in craft occupations, the percentage difference is 44 percent of the nonfarm range; for years of schooling the percentage difference is 36 percent; and for income the percentage difference is considerably less, 18 percent. When either self-employed managers or sales workers are compared with craft workers, the contrast is less noticeable, but it nonetheless remains. Comparing self-employed managers with craft workers, the difference in mean son's SEI is 30 percent of the range; the difference in mean son's years of schooling is 26 percent; and the difference in the mean son's income is 21 percent. The difference between the means for sales workers and clerical workers categories of origin as a percent of the nonfarm range is 43 percent for SEI, 33 percent for education, and 24 percent for income.

If, for example, the difference in mean income for sons of sales workers and sons of craft workers appears small, it should be remembered that the difference in the income means for the fathers probably has the opposite sign. The OCG data set does not contain any information on income of fathers, but the 1940 census data show that the median wage or salary income for year-round full-time male workers was 7.8 percent higher for the craft category than for the sales category.

In both table 5.1 and table 5.4, the pattern of the coefficients is well summarized by the farm-nonfarm distinction, the manual-nonmanual distinction, and by distinguishing the self-employed professional category. This specification aggregates the twelve major census occupations into four categories of origin. Table 5.5 presents the least-squares estimates for model 5.2a with the four-segment specification of origins. The value of $\bar{R}^2$ is only 1 percentage point below its value for the linear regression of son's SEI on father's SEI.

Constant-β segmentation model. – Let us now replace the assumption that all β_k are zero with the less restrictive assumption that all β_k have the same value. The segmentation model for son's expected SEI can now be written:

TABLE 5.4. Effect of father's major census occupation on mean income and mean education

		$\hat{\alpha}_k$		$SE(\hat{\alpha}_k)$	
k	*Category of origin*	*Income (in $)*	*Education (years)*	*Income (in $)*	*Education (years)*
1	Self-employed professionals	2,486	4.12	790	.59
2	Salaried professionals	2,913	3.60	823	.47
3	Salaried managers	2,319	2.82	582	.43
4	Self-employed managers	1,309	1.65	539	.40
5	Sales workers	1,419	2.01	570	.42
6	Clerical workers	1,186	2.16	593	.44
7	Craft workers	555	0.33	497	.37
8	Operatives	135	−0.35	300	.37
9	Service workers	16	−0.09	349	.43
10	Nonfarm laborers	− 625	−1.02	328	.41
11	Farmers	−1,174	−1.48	195	.36
12	Farm laborers	−1,301	−2.77	188	.35

	Income	*Education*
$E(\hat{Y})$	$5,588	11.29 years
$\bar{R}^2$	12.91%	20.21%

TABLE 5.5. Effect of father's segment on mean SEI

k	*Category of origin*	$\hat{\alpha}_k$	$SE(\hat{\alpha}_k)$
1	Self-employed professionals	27.0	3.1
2	Other white-collar workers	15.7	0.8
3	Blue-collar workers	− 0.9	0.5
4	Farm	−11.0	0.8

SEE = 22.60 R^2 = 16.25%

$$E(Y|X;k) = E(Y) + \alpha_k + \beta[X - E(X)] \qquad [5.2b]$$

where

k = index for father's major census occupation
X = father's SEI
Y = son's SEI.

In this model both father's major census occupation and father's SEI affect son's SEI. The least-squares fit produces an 11 percent increase in $\bar{R}^2$ over its value when the β_k are assumed to be zero.

The estimates shown in table 5.6 indicate an effect of father's occupation not picked up by father's SEI: even though their fathers are in detailed census occupations with the same SEI, two individuals may differ in expected SEI because their fathers have different broadly defined occupations. Furthermore, the estimates for α_k provide evidence that the polar version of the segmentation model is inaccurate when the categories of origin are defined by father's major census occupation: despite their fathers being in the same census occupation, two sons may differ in expected SEI if their fathers have different detailed census occupations.

TABLE 5.6. Estimates for model 5.2b

k	*Category of origin*	α_k	$SE(\alpha_k)$
1	Self-employed professionals	8.33	4.33
2	Salaried professionals	6.01	3.48
3	Salaried managers	3.76	3.25
4	Self-employed managers	4.74	2.71
5	Sales workers	7.58	2.93
6	Clerical workers	9.09	3.01
7	Craft workers	0.20	2.31
8	Operatives	1.75	2.20
9	Service workers	3.34	2.57
10	Nonfarm laborers	0.86	2.39
11	Farmers	–6.20	2.14
12	Farm laborers	–8.55	2.07

$\beta = .340$; $E(\hat{Y}) = 39.35$; $\bar{R}^2 = 18.87\%$; $SE(\hat{\beta}) = .033$; SEE = 22.24

As a further test for the effect of major census occupation controlling for SEI, fathers who were either clerical workers or craft workers were both divided into three groups according to whether their SEI scores were below 40, from 40 to 50, or above 50. The proportion of sons with a nonmanual job was compared. In every group a considerably higher percentage of clerical workers' sons had nonmanual jobs. For the low SEI groups, 73 percent of clerical workers' sons became nonmanual workers versus 46 percent of craft workers' sons; for the medium SEI groups, 65 percent versus 42 percent; and for the high SEI groups, 82 percent versus 47 percent.

Comparing the estimates of model 5.2b with the estimates of model 5.2a, father's SEI reduces the magnitude by which the nonmanual coefficients exceed the manual coefficients, but the manual-nonmanual division remains. The estimated net contribution of father's occupation (net of SEI) on son's SEI is greater for each of the nonmanual categories than for any of the manual categories.

There is less variation among the model 5.2b estimates than among the model 5.2a estimates; there remains, however, a similar pattern to the estimates. The self-employed professionals coefficient does not stand out as clearly from the other nonmanual coefficients, nor does the nonfarm laborers coefficient seem especially low compared with the other manual coefficients. Nevertheless, unless the standard errors are interpreted as too high to permit even tentative inference, there is some evidence that, even after father's SEI is taken into account, there is an isolated top and bottom to the nonfarm labor force. Compared with the two managerial origins, the two professional origins (especially self-employed professionals) have high coefficients; the assumption of a linear relationship between father's and son's SEI, combined with the assumption of a constant added advantage for all nonmanual origins, underestimates the expected SEI of the sons of professionals. Similarly, nonfarm laborers are lower in standing relative to service workers and operatives than indicated by SEI. It is also possible to detect some additional evidence for the manual-nonmanual distinction. Craft workers are unexpectedly unlike clerical workers and sales workers; instead, craft workers are more similar than expected to operatives and service workers, and clerical workers and sales workers are unexpectedly similar to managers.

Unconstrained segmentation model.– Let us now examine the fit of model 5.2 where no restrictions are imposed on the parameters:

$$E(Y|X;k) = E(Y) + \alpha_k + \beta_k[X - E(X)] . \qquad [5.2]$$

Table 5.7 presents the least-squares estimates. By allowing the effect of father's SEI to depend on father's major census occupation, model 5.2 allows

TABLE 5.7. Estimates for model 5.2

k	*Category of origin*	$\hat{\alpha}_k$	$\hat{\beta}_k$	$SE(\hat{\alpha}_k)$	$SE(\hat{\beta}_k)$
1	Self-employed professionals	20.1	.13	16.9	.20
2	Salaried professionals	10.5	.24	11.4	.14
3	Salaried managers	− 0.9	.46	9.2	.12
4	Self-employed managers	7.1	.22	6.2	.09
5	Sales workers	7.6	.34	7.2	.11
6	Clerical workers	7.7	.41	9.9	.18
7	Craft workers	0.1	.36	4.5	.06
8	Operatives	2.4	.39	4.4	.08
9	Service workers	3.5	.35	5.3	.16
10	Nonfarm laborers	− 0.8	.26	5.2	.40
11	Farmers	−19.4	−.67	14.8	1.01
12	Farm laborers	− 8.0	.37	4.1	.43

$E(\hat{Y}) = 39.35$ $E(X) = 27.04$ $\bar{R}^2 = 18.78\%$ SEE = 22.25

us to consider an additional type of segmentation effect. Relaxing the assumption of constant β_k, however, fails to improve the fit. It is hazardous, moreover, to make inferences on the basis of differences among the coefficients since, as reflected in the high standard errors, sampling error is likely to be high. Also, differences in the α coefficients are not as directly interpretable as in restricted versions of the segmentation model. Nevertheless, despite the lack of evidence that the differences among the estimated β coefficients are not the result of sampling error, explanations for the observed differences are offered in order to indicate the issues raised by the segmental perspective.

Under model 5.2a, $\alpha_k - \alpha_{k'}$ gives the difference in expected SEI between persons with origins k and k′, without taking into account differences in father's SEI. Model 5.2b divides the differential advantage in expected SEI into two parts: a part due to father's SEI and a part due to father's major census occupation. Thus $\alpha_k - \alpha_{k'}$ could be interpreted as the differential advantage of having a father in category k as opposed to k′, after taking account of the effect of differences in father's SEI. Under model 5.2 there is no single number that gives the differential advantage, holding SEI constant, of the categories of origin. Differences in the β_k estimates, however, are slight enough that the previous generalizations concerning the net effect of categories of origin still stand. According to the estimates for model 5.2, model 5.2b α estimates understate the advantage in expected SEI for sons of low SEI professionals and overstate the advantage for low SEI salaried managers. To a lesser extent, the advantage for sons of self-employed managers is understated, and the advantage for high SEI nonfarm laborers is overstated.

Variation in the β coefficients, insofar as such variation is not the result of sampling error, might be explained in a number of ways. We shall briefly consider three approaches. One approach is to consider a measurement error model. A second approach is to consider the possibility that SEI differences have a different meaning within each broad occupational category. And a third is to relate occupational categories to the working-class/middle-class distinction.

A simple measurement error model may be obtained by assuming that father's SEI has an error component which is uncorrelated with the true value of SEI, and that the proportion of the variance in SEI due to error is unrelated to the category of origin. Under these assumptions, the attenuation due to measurement error will be greatest for those categories of origin for which variation in father's SEI is low. This explanation is consistent with the low coefficient for nonfarm laborers; next to the origin of farmer, the nonfarm laborers category of origin has the smallest variance in father's SEI. However, the three other nonfarm categories of origin with low variance in father's SEI (clerical workers, operatives, and service workers) do not have low estimates for β. Moreover, the category of origin with the largest variance, self-employed professionals, has the lowest nonfarm coefficient. Attenuation due to measurement error thus does not explain the pattern of the β estimates.

For one major census occupation the variation across detailed occupations in SEI may be largely due to income differences; for another major occupation the variation may be primarily the result of educational differences among the detailed occupations. Since educational level (more exactly, the education level typical of the workers in the detailed occupation) and income level influence the father, and thereby the son, in a different way, differences in the β coefficients may be expected. Further research into this hypothesis would require measuring separately the income level and the education level of the detailed occupations of the fathers.

Despite variation from one narrow occupation to another in typical levels of income and education, for some major occupations the job character may be especially similar. For example, it could be argued that the job experience of self-employed professionals is more constant than indicated by the variations in their incomes and education, and thus the low coefficient for this origin is observed. It could be argued as well that the high coefficients for salaried managers and sales workers are due to exceptional variation in the character of their work from one detailed census category to another. For example, there is likely to be a major difference in the level of autonomy between the large number of managers and sales workers in small retail outlets, as compared to managers and sales workers in capital-intensive industries. This conjecture is consistent with the high coefficients observed for these two origins.

An alternative approach is to consider the relation of the major census occupations to broad social divisions. The manual-nonmanual distinction may be viewed as corresponding to a division of communities, schools, and both formal and informal associations into middle-class versus working-class. The well-paid, blue-collar worker, for example, may be able to move into a middle-class community; the poorly paid, white-collar worker may end up in a working-class neighborhood. On a more aggregative level, low SEI white-collar occupations may be dominated by ties to the working class. From this same perspective, virtually all professionals are securely middle class. Thus, we would expect differences in SEI to be more important for the upper manual and lower nonmanual categories. This expectation has some support in the observed coefficients: the professional and nonfarm laborer categories of origin have lower coefficients than the other manual, clerical, and sales categories of origin.

Effect of origins on the SEI variance

There are three conceptual steps in our approach to measuring the effect of origins on destinations: (1) persons are divided into categories of origin based on the socioeconomic positions of their fathers; (2) for each category of origin a distribution of destinations is determined; (3) to reveal the effect of socioeconomic origin, the distributions of destinations are compared. In contrast, the individualistic approach to measuring the effect of origins looks at a person's socioeconomic position as the outcome of adding up the effect of a number of factors, one of which is socioeconomic origins.

The previous section focused on the location of the distributions of destinations, specifically on son's expected SEI for each category of origin. The subject of this section is the dispersion of the distributions of destinations, in particular the variance in son's SEI. In the previous section, the categories of origin were broad aggregates defined by father's major census occupation. A much more detailed occupational classification scheme will now be used.

Socioeconomic category of origin is determined by father's job category. Each job category defines a different category of origin. There is a job category for each of the combinations of three categorical variables—three-digit occupation, three-digit industry, and class of employer (private, government, self-employed). This classification scheme is more detailed than the scheme underlying SEI. (Each of 446 job categories is given an SEI score; 270 of the categories are three-digit occupations, and the other 176 are industrial and class-of-worker subdivisions of three-digit occupations.) In the sample, 747 fathers were each the sole occupant of a job category, and the remaining 4,068 fathers were distributed across 474 categories.

Data analysis.– The simple linear regression of son's SEI on father's SEI is a useful point of departure. The standard assumption of a homoscedastic error term, which is implied by bivariate normality, allows us to talk about the variance in socioeconomic position for a group of persons with the same socioeconomic origins without specifying any further characteristics of the group, such as the level of the socioeconomic standing of the group. Thus, the linear regression estimator for the variance in SEI holding origins constant is

$$S^2 = \sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2 / (N-2),$$

where Y_i is the fitted value corresponding to X_i. S^2 is the unbiased estimate of the conditional variance of son's SEI, Y, given father's SEI, X. The computed value is

$$S^2 = 508.$$

The estimated variance for a group of persons randomly selected without holding origins constant is

$$S_Y^2 = \Sigma(Y_i - \bar{Y})^2 / (N-1).$$

S_Y^2 is the unbiased estimate for the unconditional variance in son's SEI. The computed value is

$$S_Y^2 = 610.$$

$\bar{R}^2$ compares the conditional variance to the unconditional variance: $\bar{R}^2$ gives the proportional difference between these two variances, using the unbiased estimates. For our data,

$$\bar{R}^2 = \frac{S_Y^2 - S^2}{S_Y^2} = 16.7\%.$$

Thus, assuming homoscedasticity, the value of $\bar{R}^2$ indicates that the dispersion for a group of persons with the same origins is 16.7 percent less than the dispersion for a random group of persons.[3]

Using detailed categories of origin, estimates of S^2 and $\bar{R}^2$ may be obtained with less restrictive assumptions as to the relationship between the socioeconomic positions of father and son. Under the simple linear regression approach it is assumed that (a) all persons with fathers in the same job category have identical expected SEI, and (b) son's expected SEI is a linear function of the SEI of the father's job category. Assumption (a) is now weakened by using a job classification scheme that constitutes a disaggregation of the 446-category job classification scheme underlying the SEI index. We do not make assumption (b): father's SEI and the assumption of linearity are not used. Although the following estimates do not depend on assumptions of linearity and the use of SEI to measure father's socioeconomic position, they do, as in the case of the estimates based on the simple linear regression, depend on assuming that all persons placed in the same category of origin have identical socioeconomic origins, that son's SEI measures son's socioeconomic position without error, and that all the distributions of destinations have the same variance and shape.

Assuming homoscedasticity, the pooled unbiased estimate of the conditional variance of son's SEI, holding constant father's job category, is

$$S^2 = \sum_{i=1}^{K} \sum_{j=1}^{N_i} (Y_{ij} - \bar{Y}_i)^2 / (N - K) = 481.,$$

where $i = 1, \ldots, K$ identifies father's job category and $j = 1, \ldots, N_i$ distinguishes sons with fathers in the same job category. (The use of detailed job categories to estimate S^2 can be expressed as the analysis of variance model, $Y_{ij} = a_i + e_{ij}$. The regression model adds the restriction $a_i = \alpha + \beta X_i$.) This estimate of the conditional variance implies a 26 percent upward revision of $\bar{R}^2$ to 21.1 percent. The pooled maximum likelihood estimate of S^2 is 358.8,

3. Instead of comparing the unbiased estimates, R^2 (the coefficient of determination) compares the maximum likelihood estimates. The unbiased estimates take into account the number of groups, K. As indicated by the following identity, $\bar{R}^2$ is always smaller in value than R^2 and therefore provides a more conservative measure of strength of relationship:

$$\bar{R}^2 = R^2 - \frac{K-1}{N-K}(1 - R^2)$$

which implies a more than doubling of $\bar{R}^2$. However, this estimate of S^2 is biased.

The division of fathers into 474 categories yields 474 unbiased estimates of the conditional variance of son's SEI holding constant socioeconomic origins.

$$S_i^2 = \sum_j (Y_{ij} - \bar{Y})^2 / (N_i - 1) \quad (i = 1, \ldots, 474).$$

The variation in these estimates is shown by table 5.8. The median estimate is 384. If the median estimate were substituted for S^2 in the $\bar{R}^2$ formula, the value of $\bar{R}^2$ would increase from 16.7 to 37.1 percent:

$$\bar{R}^2 = \frac{S_Y^2 - S^2}{S_Y^2} = 37.1\%,$$

where S^2 is the estimated variance for son's SEI assuming homoscedasticity, and S_Y^2 is the unconditional variance.

TABLE 5.8. Conditional variance for son's SEI

Range of values	*Number of values*	*Cumulative percentage*
0 – 200	141	30
201 – 400	100	51
401 – 600	89	70
Above 600	144	100

The main argument in support of the median estimate rather than the pooled estimate is that there are numerous reasons for suspecting that fathers placed in the same job category are not identical in socioeconomic standing. High estimated conditional variance may be due to an exceptionally ill-defined or heterogeneous category. The median estimate is more robust than the pooled estimate to such classification error.

Rather than regard the heteroscedasticity as spurious in the sense of merely reflecting the differential homogeneity of job categories, heteroscedasticity may be interpreted as the result of differences among origins in the extent to which socioeconomic origins influence socioeconomic destinations. In this regard the median estimate may be regarded as an indication of what is typical.

Returning to the theme that it is the top and the bottom of the socioeconomic structure that merit particular attention, let us estimate S^2 separately for those with high and low socioeconomic origins. Table 5.9 provides evidence that the dispersion of son's SEI is relatively low for the top and the bottom.

Observations were divided into three strata on the basis of father's SEI, roughly a bottom 60 percent, a middle 30 percent, and a top 10 percent. (Of the 4,815 fathers in the sample, 58.6 percent of the fathers have SEI scores of 20 or less, 30.0 percent have SEI scores from 21 to 59, and 11.4 percent have SEI scores of 60 or above.) Assuming homoscedasticity for detailed job categories in the same stratum, pooled unbiased estimates for the conditional variance of son's SEI were calculated for each of the three strata. Compared to the implications of the simple correlation, the variance is higher than anticipated for the middle stratum, but lower for the top and bottom strata. The implied $\bar{R}^2$ statistics may be compared to the estimate based on the correlation of father's and son's SEI: $\bar{R}2$ for the bottom stratum is increased

TABLE 5.9. Effect of origins (defined by interval of father's SEI) on SEI variance

k	*Father's SEI*	N_k	K_k	S_k^2	R_k^2	$R_k^2/\bar{R}^2$
1	1–20	2583	167	459.7	24.6%	1.47
2	21–59	1096	210	536.2	12.1	.72
3	60–96	389	97	486.6	20.2	1.21

k = category of origin.
N_k = number of persons with origin k.
K_k = number of detailed job categories across which persons of origin k are distributed.
S_k^2 = pooled unbiased estimate of SEI variance for those with origin *k*.
$\bar{R}_k^2$ = $1 - S_k^2/S_Y^2$, where S_Y^2, the unconditional variance of SEI, equals 609.75 (calculations based on S_k^2 to two decimal places).
$\bar{R}^2$ = $1 - S^2/S_Y^2$, where S^2 is the unbiased estimate of SEI variance holding origins constant from the regression of son's SEI on father's SEI ($\bar{R}^2$ = 16.73%).

by about 50 percent, and for the top stratum by about 20 percent. Thus it is much more difficult to predict a person's socioeconomic position if his father is in the middle stratum than if his father is in the top or bottom stratum.

Let us now use the detailed job categories to estimate the variance in son's SEI for categories of origin defined by father's major census occupation: we pool the estimates of conditional variances for all detailed categories which are subdivisions of a major category. For example, it is assumed that sons of doctors have the same SEI variance as do sons of lawyers, but it is not assumed that expected SEI is the same for sons of doctors and sons of lawyers. The estimates for the variances for each of the twelve broad categories of origin are presented in table 5.10, with the origins ranked and grouped by the size of the variance.

The SEI variance is especially low for sons of farm laborers. Also low in SEI variance are sons of farmers, self-employed professionals, and operatives. No occupation has a uniquely high variance; instead, four categories of origin are characterized by high dispersion in son's SEI. The high-variance occupations, consisting of two manual and two nonmanual categories of origin, are nonfarm laborers, service workers, and the two managerial categories.

The pattern of SEI variances by father's major census occupation requires that we revise the previous conclusion based on three SEI intervals that SEI variance is highest for those with middle-level origins and lower for those whose fathers were either in the top range or bottom range of the occupational structure. As shown in table 5.10, the origins which are at the middle in variance are also at the middle of the occupational structure: the top manual category (craft workers), and the three nonmanagerial, nonself-employed nanmanual categories (salaried professionals, sales workers, and clerical workers). The low dispersion for the top and bottom interval of father's SEI may be attributed to the self-employed professionals category and to a subset of the low-standing categories of origin, namely farm origins (especially farm laborers) and operatives; two low-standing origins, nonfarm laborers and service workers, are high in dispersion of son's SEI.

Rather than all self-employed professionals, let us consider all self-employed professionals with SEI above 90. This category of origin consists of all fathers who were self-employed architects, dentists, physicians and surgeons, and lawyers. For this category of origin, the pooled unbiased estimate of the variance in son's SEI is 394. The implied reduction in variance is 35.4 percent, which is more than twice the dispersion reduction that would be expected from the fitted regression of son's SEI on father's SEI.

TABLE 5.10. Effect of origins on SEI variance

k	*Father's major census occupation*	S_k^2	$\bar{R}_k^2$	$\bar{R}_k^2/\bar{R}^2$
12	Farm laborers	280	54.0	3.23
11	Farmers	457	25.1	1.50
1	Self-employed professionals	457	25.0	1.50
8	Operatives	470	22.9	1.37
2	Salaried professionals	501	17.8	1.07
6	Clerical workers	504	17.3	1.04
7	Craft workers	507	16.8	1.01
5	Sales workers	510	16.3	0.98
10	Nonfarm laborers	540	11.4	0.68
4	Self-employed managers	541	11.3	0.68
9	Service workers	545	10.6	0.63
3	Salaried managers	552	9.4	0.56

S_k^2 = pooled unbiased estimate of SEI variance for those with origin k.

$\bar{R}_k^2$ = $1 - S_k^2/S_Y^2$, where S_Y^2, the unconditional variance of SEI, equals 609.75 (calculations based on S_k^2 to several decimal places).

$\bar{R}^2$ = $1 - S^2/S_Y^2$, where S^2 is the unbiased estimate of SEI variance holding origins constant from the regression of son's SEI on father's SEI.

CHAPTER SIX

CONCLUSIONS

THIS STUDY analyzed data on intergenerational mobility by applying several analytical techniques which embody a segmentation perspective. The results are reviewed in the first two sections of this chapter. The first section is concerned with the socioeconomic standing of broadly defined occupations; and the second section examines the magnitude of the impact of socioeconomic origins. A final section discusses further application of the techniques developed in this study.

Relative socioeconomic standing of occupations

It is a popular view that the American socioeconomic structure is characterized by a vast undifferentiated middle class. The data on intergenerational mobility contradict this view. The importance of the manual-nonmanual distinction is clearly apparent in the intergenerational mobility table discussed in chapter three. In particular, the probability of being in the top category is very low for persons with manual origins, while the probability of being a laborer is low for those with nonmanual origins.

If the American socioeconomic structure were well characterized by a vast middle class, we would expect relatively short "mobility distance" between occupations which differ on the basis of the manual-nonmanual distinction. The cluster analysis of the major census occupations presented in chapter four, which used a variety of definitions of mobility distance, consistently found a manual-nonmanual division in the occupational structure. On the basis of downward mobility or occupational prospects, the division is clearest. On the basis of social composition, it is ambiguous whether the clerical occupation belongs to the nonmanual cluster, whereas on the basis of upward mobility there is a question whether the craft occupation belongs to the manual cluster. There is a tendency for these two occupations on the border each to be singleton clusters.

The second important division in the occupational structure is the farm boundary. This feature is most clearly revealed by differences in social composition: few with nonfarm origins have farm destinations. However, the probabilities of upward mobility from farm origins are not strikingly dissimilar to the rates of upward mobility for other manual workers, especially service workers and laborers.

The social isolation of the very top occupation is the third salient feature. This feature is most evident on the basis of occupational prospects and least evident on the basis of direct mobility flows. There is not much support for grouping together supervisory and professional workers.

A final feature of the occupational structure, but one which is not clearly distinguishable by the data used, is the isolation of a bottom category of workers. An isolated (nonfarm) bottom category is not revealed by differences in occupational prospects or social composition, but it is evident in upward and downward mobility flows. There is support, however, for several different definitions of the bottom category: (1) laborers only, (2) laborers and service workers, (3) laborers, service workers, and operatives.

Importance of socioeconomic origins

The 12-by-12 mobility table, based on a classification of males by their fathers' and their own major census occupations, provides clear evidence that a person's socioeconomic destination is affected by his socioeconomic origin. A person's probability of being in any occupation is sensitive to the similarity of that occupation to his father's occupation.

Strict inheritance. – With the minor exception that the probability of being an operative is slightly higher for sons of nonfarm laborers than for sons of operatives, the probability of entry into any of the occupations is greatest for sons whose fathers are in the same occupation. For example, the probability of being a salaried professional is greater than one in three for sons of salaried professionals. Grouping together all those whose fathers were not salaried professionals, the probability of becoming a salaried professional is about one in ten. Taking the ratio of these two probabilities, the computed odds favoring strict inheritance are 3.2. This value is typical, though there is a great deal of variability in this probability ratio, ranging from a low of 1.5 for operatives and clerks to 13.0 for self-employed professionals and 14.5 for farmers.

The odds favoring inheritance, on the average for the twelve occupations, are 4.5. The weighted average, determined by weighting by the proportion of sons in each occupation, is 2.7. Weighting by the proportion of fathers in each occupation, the average is 5.8. Thus a conservative summary is that, on average, entry into an occupation is three times more likely for sons with fathers in that occupation.

Strict inheritance implied by SEI correlation.– Using the often made normality assumption, the odds favoring inheritance may be theoretically derived from the father-son correlation. The derived values and actual values may be compared to further clarify that the correlation approach tends to understate the effect of origins on destinations.

Let us assume that the son's SEI and father's SEI are distributed in accordance with the bivariate normal probability distribution. Let us define the categories of origin and categories of destination to be tenths or deciles of the SEI distributions of categories of origins and destinations. Thus the transition probability for the cell in the first row and first column of the theoretically derived contingency table is the derived probability that the son's SEI is in the top 10 percent of all sons' given that the father's SEI is in the top 10 percent of all fathers'. We shall assume, in accordance with the results of Blau and Duncan (1965:6), that the value of the father-son SEI correlation coefficient is 0.38.[1]

The derived average odds favoring strict inheritance are 1.6. In contrast, based on the twelve major census categories, depending on how the twelve values are averaged, the odds favoring strict inheritance are either 2.7, 4.5, or 5.8. For the derived table, six of the ten deciles have odds favoring strict inheritance no higher than 1.2. In contrast, none of the twelve major census occupations have odds favoring strict inheritance below 1.5. The highest value of the inheritance odds for the implied table is 2.8, whereas five of the twelve values for the directly observed table are at least as great as 3.2.

The failure of the correlation model to account for the observed degree of strict occupational inheritance does not appear to be correctable by using either broader or narrower percentile intervals. To investigate the effect of the size of the categories of origin and destination, the probability of strict

1. The decile transition probabilities were approximated by conditioning on the midpoint of the decile intervals. For example, the transition probability for the cell in row 1 and column 1 is conditional on father's SEI being at the ninety-fifth percentile rather than on father's SEI being between the ninetieth and one hundredth percentiles. The transition probabilities were found using the following formula for the probability Y exceeds the 100a

occupational inheritance was computed for the second quintile and for the interval covering the percentiles interval from 27.5 to 32.5. The implied odds favoring strict inheritance were found to be 1.22 for the second quintile and 1.18 for the percentiles interval from 27.5 to 32.5.

Top and bottom destinations. – To investigate the extent to which access to a privileged occupation (high material rewards, prestige, and autonomy) is sensitive to origins, particular attention was given to the probabilities of being a self-employed professional. In parallel, to obtain information on the risk of an unfortunate occupation, attention was focused on the probabilities of being a nonfarm laborer.

The probability of being a self-employed professional is low for men whose fathers were not self-employed professionals. The chances for sons of self-employed professionals relative to the chances for the sons of fathers from each of the other nonmanual categories of origin are about 5 to 1. Relative to the chances for sons of nonfarm laborers and sons of farm laborers, the chances for sons of self-employed professionals are 65 to 1 and 84 to 1. Relative to the chances for sons of fathers from other manual categories of origin, the chances are about 20 to 1.

The risk of being a nonfarm laborer is much more equally shared than is

---

percentile given that X is at its 100b percentile, where Y_a is the 100a percentile of the marginal distribution of Y, and X_a is the 100b percentile of the marginal distribution of X:

$$\text{probability } (Y \leqslant Y_a) = a;$$
$$\text{probability } (X \leqslant X_b) = b.$$

Given that X and Y are joint normal with correlation ρ, then

$$\text{probability } (Y \leqslant Y_a | X_b) = 0.5 - 0.5P(\gamma),$$

where

$$\gamma = \frac{P^{-1}(2a-1) - \rho P^{-1}(2b-1)}{\sqrt{1-\rho^2}};$$

$$P(\gamma) = \frac{1}{\sqrt{2\pi}} \int_{-\gamma}^{\gamma} e^{-1/2} z^2 \, dz.$$

Note that $P^{-1}(\ )$ is the inverse function of $P(\ )$.

the likelihood of being a self-employed professional. Occupational origins, however, remain important. Sons of nonfarm laborers and sons of farm laborers have virtually the same probability of being a laborer, which is about twice the probability for sons with fathers from other manual categories of origin and about seven times the probability for sons with fathers from nonmanual origins.

Combining a continuum model and a segmental perspective.– Chapter 5 applied a continuum view of socioeconomic destinations. Rather than define categories of destination arbitrarily, categories of destination were defined by a quantitative measure of socioeconomic position, in particular Duncan's SEI index. Following the dispersion reduction approach, the errors in the fit for individual observations (the residuals) were relied upon to indicate the strength of relationship. For example, using the polar segmentation model, the predicted SEI for a welder's son is the mean SEI for all sons of craftsmen. Using the regression model, the predicted SEI for a welder's son is found by substituting the welder's SEI into the fitted linear regression equation of father's SEI on son's SEI. $\bar{R}^2$ for both models is determined by the squared errors in predicted SEI.

Even though the polar segmentation gives only a slightly higher $\bar{R}^2$, the performance of this model using the major census occupations is an indication that $\bar{R}^2$ is in need of upward revision. The reason for upward revision is the obvious distortion in the assumption of identical socioeconomic position for all fathers in the same major census occupation. The addition of father's SEI, with and without the assumption of a constant coefficient for the twelve categories of origin, is a step toward taking into account the socioeconomic heterogeneity among fathers in the same major census occupation. There was an improvement in fit but by no means was the improvement dramatic; the proportional increase in $\bar{R}^2$ is above 10 percent. There remains, however, considerable reason to doubt the adequacy of SEI to capture the socioeconomic differences among fathers and thus it can be argued that measurement error in SEI has a major downward effect on $\bar{R}^2$.

Based on a division of persons by detailed occupational category, $\bar{R}^2$ was determined without using father's SEI. It was assumed that persons in the same detailed category of origin have the same expected SEI and that for all categories of origin the variance in sons' SEI is constant. The proportional increase in $\bar{R}^2$ over the simple regression model is more than a quarter.

There appears to be considerable heteroscedasticity across the detailed categories of origin in son's SEI. The median variance in son's SEI is 37 percent less than the overall variance in son's SEI. The variance in son's SEI also is below average for the higher categories of origin and the lower categories of origin, suggesting that much of the observed mobility is due to movement of sons who start out in the middle as opposed to those who start at the top or the bottom.

Future research

This study has relied on the twelve major categories of occupations used by the Census Bureau and the distinction between manual and nonmanual occupations. Use of these categories reflects the decision first to discover what can be done with the mobility data using the conventional categories. Not only are these categories convenient; results produced by use of these categories also are likely to receive less resistance than results depending on the introduction of new categories. An obvious future task is to repeat the analysis using improved definitions of labor segments.

An important direction in which to extend this analysis is to investigate intragenerational mobility. Many of the same methods may be applied to intragenerational mobility in order to determine differences and relationships among labor segments and to define labor segments that reflect socioeconomic structure.

Application of the segmentation perspective to the OCG data on intergenerational mobility provides a stronger sense of socioeconomic inequality than does previous work based on the individualistic perspective. The results may well be considerably strengthened by better data, by more robust methods of estimation, and by improved definitions of labor segments.

AFTERWORD

Chapter 1 of the text criticizes *Inequality,* a widely noted book by Christopher Jencks and associates. After my text was written, Jencks and associates published a new book, *Who Gets Ahead?,*[1] which appears to argue for the importance of socioeconomic origins, in an apparent reversal of the position taken in *Inequality.*

Inequality argued that social origins were not very important, stating that "the role of a father's family background[2] in determining his son's status is surprisingly small, at least compared to most people's preconceptions" (p. 179). *Who Gets Ahead?* appears to provide a radically revised assessment of the role of origins. For example, *Who Gets Ahead?* concludes that "[family] background exerts a larger influence on economic outcomes than past research has suggested. . . ." (p. 229).

1. Christopher Jencks, S. Bartlett, M. Corcoran, J. Crouse, D. Eaglesfield, G. Jackson, K. McClelland, P. Mueser, M. Olneck, J. Schwartz, S. Ward, and J. Williams, *Who Gets Ahead?: The Determinants of Economic Success in America* (New York: Basic Books, Inc., Publishers, 1979). Chapter 1 of *Intergenerational Occupational Mobility* was made available to Jencks prior to the publication of *Who Gets Ahead?*. For another perspective, see Peter Meyer, "The Reproduction of the Distribution of Income," Ph.D. diss., University of California, Berkeley, 1979.

2. From the text, it is clear that "father's family background" refers to the family background of the son, not the family background of the father.

Unfortunately, *Who Gets Ahead?* does not explain this change in position. Since a primary objective of the new book was to use better data (see chapter 1 of *Who Gets Ahead?*), the reader might falsely assume that the change in position is due to new data. The new position is instead due to differences in concepts and changes in method of statistical interpretation.

Inequality focused on the correlation between the socioeconomic positions of father and son. In contrast, *Who Gets Ahead?* focuses on the correlation between the socioeconomic positions of brothers. The shift from father-son correlation to brothers correlation constitutes a subtle but important change, which is unfortunately not clarified by Jencks.

The magnitude of the father-son correlation is a measure of the degree to which socioeconomic position in one generation is passed to the next generation. It is a measure of the degree to which privilege is monopolized by favored social strata.

Correlation in the socioeconomic positions of brothers may be attributed to a number of factors other than socioeconomic position of father. Among these factors are genetic characteristics and child-centeredness of parents. Thus the brothers correlation, unlike the father-son correlation, does not directly measure the persistence of class position. Father-son correlation measures amount of social (im)mobility, brothers correlation does not.

In addition to turning away from the concept of social mobility, *Who Gets Ahead?* differs from *Inequality* in its method of interpreting a correlation coefficient.

Inequality depended heavily on a reduction-in-inequality interpretation of statistical correlation, which I criticized in chapter one of this book as seriously misleading. Applied to the correlation in the socioeconomic positions of brothers, the reduction-in-inequality method asks how much inequality there is between two brothers as compared to the amount of inequality between two men picked at random. The impression created by this method, whether using data from *Inequality* or *Who Gets Ahead?*, is that family background is not very consequential.

The regression-slope method, which I recommended in chapter 1 of this book, is employed in *Who Gets Ahead?* (p. 214). Applied to the brothers correlation, this method asks what is the advantage in socioeconomic position typically accruing to a son who has a certain amount of advantage in family background. Using either the new data or the old data, this method of interpretation shows that family background has a very substantial impact.

In sum, *Who Gets Ahead?* is a major improvement over the earlier book in that the authors are not misled by incautious interpretation of statistical correlation to conclude that socioeconomic origins are not very important. However, since the focus is shifted from socioeconomic origins to family background, *Who Gets Ahead?* does not measure the amount of social mobility, nor does it directly address the importance of socioeconomic origins.

Who Gets Ahead? describes and explains social inequality using individuals rather than groups as the units of analysis, and in this regard is consistent with *Inequality* and most other contemporary quantitative research in social stratification,[3] but contrasts with my orientation. The use of the individualistic perspective in *Who Gets Ahead?* , and elsewhere, does not characterize social inequality in a readily comprehensible way, directs attention toward competition among individuals and away from conflict among groups, and points to individual traits rather than social forces as the causes of inequality.

3. For example, O. D. Duncan, D. Featherman, and B. Duncan, *Socioeconomic Background and Achievement* (New York: Seminar Press, 1972), and W. H. Sewall and R. M. Hauser, *Education, Occupation, and Earnings* (New York: Academic Press, 1975).

APPENDICES

APPENDIX A

MOBILITY TABLES

TABLE A.1. Matrix of transition probabilities, *T*

	Destination												
Origin	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*Total*
1	17.55	33.40	10.36	4.65	11.42	4.23	6.34	4.44	1.90	2.75	2.11	0.85	100.00
2	3.45	33.45	13.53	5.00	8.02	8.02	9.66	12.67	3.53	1.64	0.86	0.17	100.00
3	3.81	24.35	20.93	8.45	7.91	8.22	14.92	7.08	1.60	2.05	0.53	0.15	100.00
4	3.94	14.51	19.43	16.93	9.64	6.51	13.40	9.23	2.98	1.82	1.19	0.42	100.00
5	2.71	16.78	19.05	10.92	15.02	6.15	11.87	10.40	3.22	2.05	1.68	0.15	100.00
6	2.43	25.63	12.25	5.56	7.82	9.64	16.94	9.21	6.08	3.04	1.39	0.00	100.00
7	0.99	12.00	9.04	7.48	4.73	7.76	29.38	17.50	5.15	4.81	0.80	0.37	100.00
8	0.88	10.80	5.63	6.58	4.39	6.64	23.85	25.91	5.93	7.54	0.93	0.93	100.00
9	0.80	9.29	7.87	6.33	5.72	9.47	20.97	20.91	11.07	6.27	1.05	0.25	100.00
10	0.27	5.61	4.34	3.67	3.62	8.01	22.53	26.33	9.10	14.16	1.22	1.13	100.00
11	0.67	4.62	4.43	7.12	2.47	4.70	19.68	20.47	5.25	8.50	17.84	4.26	100.00
12	0.21	2.08	3.12	4.37	1.98	3.85	20.48	25.99	8.11	13.41	6.24	10.19	100.00
All	1.58	11.28	8.66	7.58	5.18	6.60	20.77	18.64	5.47	6.53	5.88	1.84	100.00

T_{ij} = conditional probability that son is in occupation *j*, given that father is in occupation *i*. Expressed as a percentage.

Key to column numbers:

1. Self-employed professional, technical and kindred workers (self-employed professionals).
2. Salaried professional, technical, and kindred workers (salaried professionals).
3. Salaried managers and administrators, except farm (salaried managers or managers).
4. Self-employed managers and administrators, except farm (self-employed managers or proprietors).
5. Sales workers.
6. Clerical and kindred workers (clerical workers).
7. Craft and kindred workers (craft workers).
8. Operatives.
9. Service workers.
10. Laborers, except farm (laborers).
11. Farmers and farm managers (farmers).
12. Farm laborers and supervisors (farm laborers).

TABLE A.2. Matrix of inflow percentages, *Y*

	Destination												
Origin	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*All*
1	15.46	4.12	1.67	0.85	3.07	0.89	0.43	0.33	0.48	0.59	0.50	0.64	1.39
2	7.45	10.12	5.34	2.25	5.29	4.15	1.59	2.32	2.21	0.86	0.50	0.32	3.41
3	9.31	8.35	9.35	4.31	5.91	4.82	2.78	1.47	1.13	1.22	0.35	0.32	3.87
4	19.74	10.17	17.75	17.66	14.72	7.81	5.10	3.92	4.31	2.21	1.60	1.76	7.91
5	6.89	5.97	8.84	5.78	11.65	3.75	2.30	2.24	2.37	1.26	1.15	0.32	4.02
6	5.21	7.70	4.79	2.48	5.12	4.95	2.76	1.67	3.77	1.58	0.80	0.00	3.39
7	11.55	19.65	19.28	18.21	16.88	21.73	26.12	17.34	17.39	13.60	2.50	3.68	18.47
8	8.57	14.77	10.03	13.39	13.08	15.53	17.72	21.45	16.75	17.79	2.45	7.84	15.43
9	2.42	3.94	4.35	4.00	5.29	6.87	4.83	5.37	9.69	4.59	0.85	0.64	4.79
10	1.12	3.24	3.26	3.14	4.55	7.90	7.06	9.19	10.82	14.10	1.35	4.00	6.51
11	11.92	11.45	14.31	26.28	13.36	19.95	26.52	30.74	26.87	36.40	84.93	64.80	27.99
12	0.37	0.52	1.02	1.63	1.08	1.65	2.79	3.95	4.20	5.81	3.00	15.68	2.83
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Y_{ij} = conditional probability that father is in occupation *i*, given that son is in destination *j*.

APPENDIX B

SOCIAL DISTANCE MATRICES

TABLE B.1. Dissimilarity in destinations (occupational prospects), $D = (d_{ij})$.

	2	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
1	.41	.59	.73	.67	.56	.96	1.04	1.01	1.21	1.21	1.32
2		.34	.49	.44	.28	.64	.76	.70	.92	.99	1.12
3			.30	.31	.26	.63	.78	.71	.96	.99	1.16
4				.20	.45	.61	.73	.70	.94	.94	1.11
5					.43	.62	.74	.70	.95	.93	1.10
6						.50	.60	.50	.76	.83	.95
7							.25	.27	.47	.55	.71
8								.23	.24	.43	.45
9									.32	.48	.55
10										.47	.30
11											.40

$d_{ij} = \sum_k |T_{ik} - T_{jk}|$, where T_{ij} = probability (destination = j|origin = i).

TABLE B.2. Dissimilarity in origins (occupational composition), *D*

	2	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
1	.46	.37	.61	.46	.75	.98	1.04	1.00	1.13	1.52	1.42
2		.30	.47	.31	.40	.63	.74	.69	.90	1.52	1.39
3			.32	.20	.43	.66	.75	.71	.91	1.45	1.32
4				.36	.36	.37	.45	.42	.60	1.20	1.07
5					.39	.63	.68	.63	.84	1.47	1.32
6						.29	.41	.33	.58	1.33	1.18
7							.25	.24	.40	1.17	1.03
8								.18	.25	1.09	.92
9									.31	1.16	.99
10										.97	.76
11											.44

$d_{ij} = \sum_k |Y_{ki} - Y_{kj}|$, where Y_{ij} = probability (origin = i | destination = j).

TABLE B.3. Upward mobility, *D*

	2	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
1	.46	.42	.40	.58	.65	1.60	1.80	1.98	5.82	2.35	7.60
2		.46	.78	.67	.44	.94	1.04	1.21	2.01	2.44	5.43
3			.45	.45	.71	.96	1.54	1.10	1.99	1.96	2.78
4				.69	1.36	1.01	1.15	1.20	2.07	1.06	1.74
5					.66	1.09	1.18	.91	1.43	2.09	2.62
6						.85	.99	.70	.82	1.40	1.72
7							.87	.99	.92	1.06	1.01
8								.89	.71	.91	.72
9									.60	1.04	.67
10										.77	.49
11											.94

d_{ij} = reciprocal of index of association corresponding to upward mobility between occupations *i* and *j*.

TABLE B.4. Downward mobility, *D*

	2	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
1	.34	.84	1.63	.45	1.56	3.27	4.20	2.87	2.38	2.78	2.18
2		.83	1.52	.65	.82	2.15	1.47	1.55	3.99	6.82	10.67
3			.90	.65	.80	1.39	2.63	3.42	3.18	11.03	12.09
4				.54	1.01	1.55	2.02	1.84	3.58	4.94	4.49
5					1.07	1.75	1.79	1.70	3.19	3.49	12.56
6						1.23	2.02	.90	2.15	4.23	21.18
7							1.06	1.06	1.36	7.38	5.02
8								.92	.87	6.29	1.97
9									1.04	5.62	7.48
10										4.81	1.63
11											.43

d_{ij} = reciprocal of index of association corresponding to downward mobility between occupations *i* and *j*.

TABLE B.5. Indexes of association

	Destination											
Origin	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
1	11.10	2.96	1.20	0.61	2.20	0.64	0.31	0.24	0.35	0.42	0.36	0.46
2	2.18	2.96	1.56	0.68	1.55	1.22	0.46	0.68	0.65	0.25	0.15	0.09
3	2.41	2.16	2.42	1.11	1.53	1.25	0.72	0.38	0.29	0.31	0.09	0.08
4	2.50	1.29	2.24	2.23	1.86	0.99	0.65	0.50	0.54	0.28	0.28	0.22
5	1.71	1.49	2.20	1.44	2.90	0.93	0.57	0.56	0.59	0.31	0.29	0.08
6	1.54	2.27	1.42	0.73	1.51	1.46	0.82	0.49	1.11	0.47	0.24	0.00
7	0.63	1.06	1.04	0.99	0.91	1.18	1.41	0.94	0.94	0.74	0.14	0.20
8	0.56	0.96	0.65	0.87	0.85	1.01	1.15	1.39	1.09	1.15	0.16	0.51
9	0.51	0.82	0.91	0.84	1.10	1.44	1.01	1.12	2.03	0.96	0.18	0.13
10	0.17	0.50	0.50	0.48	0.70	1.21	1.09	1.41	1.66	2.17	0.21	0.61
11	0.43	0.41	0.51	0.94	0.48	0.71	0.95	1.10	0.96	1.30	3.03	2.32
12	0.13	0.18	0.36	0.58	0.38	0.58	0.99	1.39	1.48	2.05	1.86	5.54

APPENDIX C

Cluster Analysis of the Laumann Data

THIS APPENDIX determines the social divisions in the labor force using a 5-by-5 intergenerational mobility table for a sample of males from a particular urban community (Laumann, 1966:79). Although this appendix supports the substantive conclusions in chapter 4, its primary function is to clarify the use of cluster analysis by treating a simpler case. The basic method used, the same as in the chapter, is to consider all cluster configurations which honor an a priori rank ordering of the points to be clustered. For a given rank ordering and a given social distance matrix, in the case of twelve occupational categories there are 2,046 alternative configurations to be compared. In the case considered in this appendix, five categories, there are but fourteen configurations to be compared.

Table C.1 presents the twenty-five rates of mobility from the five occupational origins to the five occupational destinations. The rate of mobility is the standard index of association used in the literature on occupational mobility and is equal to the ratio of observed to chance frequency. (Chance under the assumption that the origins and destinations are distributed independently: probability [origin = a, destination = b] = probability [origin = a] X probability [destination = b].) For analytical convenience, it is helpful to have an inverse measure of rate of mobility, a measure of the distance in terms of direct mobility between an origin and a destination. To construct table C.2 the elements of table C.1 are replaced by their reciprocals.

TABLE C.1. Ratios of observed frequencies to chance frequencies

Father's occupation	Son's occupation A	B	C	D	E
A	3.6	1.1	.4	.0	.0
B	1.6	2.1	1.3	.3	.1
C	.7	1.4	1.5	.3	.1
D	.5	.7	.9	1.8	1.2
E	.2	1.1	1.0	1.2	3.0

TABLE C.2. Ratios of chance frequencies to observed frequencies

Father's occupation	Son's occupation A	B	C	D	E
A	.278	.909	2.500	–	–
B	.625	.476	.769	3.333	10.000
C	1.429	.714	667	3.333	10.000
D	2.000	1.429	1.111	.556	.833
E	5.000	.909	1.000	.833	.333

The matrix of twenty-five elements can be collapsed by one of several methods into a triangular matrix of ten elements. Each element of this new matrix, to be referred to as the social distance matrix, indicates the distance between two occupations on the basis of the mobility data. For purposes of this appendix we shall consider only one specification of these elements: the social distance from origin A to destination B is the average of the mobility from A to B and from B to A. Table C.3, so computed, in terms common to the cluster analysis literature, is the dissimilarity matrix.

TABLE C.3A. Dissimilarity Matrix

	A	*B*	*C*	*D*	*E*
A	X	.74	1.82	4.0	10.0
B	.74	X	.74	2.0	1.67
C	1.82	.74	X	1.67	1.82
D	4.0	2.0	1.67	X	.83
E	10.0	1.67	1.82	.83	X

TABLE C.3. Dissimilarity matrix

	A	*B*	*C*	*D*	*E*
A	X	.77	1.96	4.0 [a]	10.0 [a]
B	.77	X	.74	2.38	5.45
C	1.96	.74	X	2.22	5.50
D	4.0 [a]	2.38	2.22	X	.83
E	10.0 [a]	5.45	5.50	.83	X

a. Found by averaging elements of table C.1 and then taking the inverse.

Key to column headings:

A – Top professional, business
B – Semiprofessional, middle business
C – Clerical, small business
D – Skilled manual
E – Semiskilled, unskilled manual

Note: Table C.3A was computed prior to table C.3 and was used in the analysis. For the most part, the differences between the two tables are minor. The elements of table C.3A are found by first converting table C.1 into a symmetric matrix and then finding the reciprocals, rather than finding the reciprocals before forming the symmetric matrix.

An examination of either the rows or columns of the social distance matrix (the dissimilarity matrix) indicates that social distance increases monotonically with movement away from the diagonal. (There is only one minor exception to this pattern: mobility between C and E is slightly more impeded than between B and E.) Thus the overall pattern verifies that the assumed rank ordering of the occupations (ABCDE) correctly reflects the social distance among the occupations, at least as currently operationalized as the average of mobility flows: mobility is highest among adjacent occupations and decreases among occupations separated by the intervening occupations. (For example, row D of the dissimilarity matrix indicates that there is more mobility between D and E [or D and C] than between D and B, and there is more between D and B than between D and A.) As a consequence, it is assured that the cluster analysis approach used, which looks only at configurations which preserve the given rank ordering, will find the optima according to the individuation criteria. No configuration is admissible which violates the given rank ordering. For example, if D and B are assigned to the same cluster, then unless C is also assigned to the cluster containing D and B, it will not be possible to improve the individuation of the clusters without changing the number of clusters.

To aid in interpreting the symbolic representation of the alternative cluster configurations in table C.4, note that a comma represents the division between clusters. The representations of the configurations corresponding to the four basic models of class structure mentioned in section 1 are these:

Isolated top	A, BCDE
Isolated bottom	ABCD, E
Undifferentiated middle	A, BCD, E
Collar-line split	ABC, DE.

It is useful to note hierarchical association among the models. For example, the undifferentiated middle model may be regarded as a combination of the features of the isolated top and the isolated bottom model. The hierarchical issue is a question of which feature is primary. Can, for example, the undifferentiated middle model be regarded as the result of dividing cluster BCDE into two clusters after having previously established that the most important single division is between A and B?

Cluster configurations involving different numbers of clusters are difficult to compare. Increasing the number of clusters will always improve cluster compactness, making it difficult to compare individuation indexes. In addition, clustering, as a form of data reduction, entails ignoring

TABLE C.4. Individuation for all configurations

Cluster configuration	*Cluster separation Distance between clusters*		*Size of clusters Distance within clusters*		*Index of individuation Ratio of separation to size*	
	Mean (1)	*Gap (min) (2)*	*Mean (3)*	*Gap (max) (4)*	*Mean (5)*	*Gap (6)*
A, BCDE	4.1	.7	1.5	1.7	2.7	.4
AB, CDE	3.4	.7	1.3	1.7	2.6	.4
ABC, DE	3.5	1.7	1.0	.8	3.5	2.1
ABCD, E	3.5	.8	1.8	1.7	1.9	.5
A, B, CDE	3.0	.7	1.4	1.7	2.1	.4
A, BC, DE	3.0	.7	.8	.8	3.8	.9
A, BCD, E	3.0	.7	1.2	1.7	2.5	.4
AB, C, DE	3.0	.7	.8	.8	3.8	.9
AB, CD, E	2.9	.7	1.2	1.7	2.4	.4
ABC, D, E	3.1	.8	1.1	.7	2.8	1.1
AB, C, D, E	2.7	.7	.7	.7	3.9	1.0
A, BC, D, E	2.7	.7	.7	.7	3.9	1.0
A, B, CD, D	3.6	.7	1.7	1.7	1.5	.4
A, B, C, DE	2.7	.7	.8	.8	3.4	.9

Note: Look for large values in columns (5) and (6) to indicate the configurations for which the clusters are better individuated. Also, look for large values in columns (1) and (2) and small values in columns (3) and (4).

information; the capacity to absorb detail may be enhanced by beginning with a small number of clusters and then subsequently examining configurations entailing a larger number of clusters. This process is likely to be the most illuminating one if the clusters are hierarchically related.

On the basis of the gap individuation index, the collar line model is clearly best. The value of the gap index is 2.1, indicating that the gap between the two clusters is 2.1 times the largest within-cluster gap. For the thirteen other configurations, this index never goes above 1.1. The explanation is that the distance between occupations C and D is at least twice the distance between any other adjacent occupations.

The mean-distance individuation index also indicates that the collar line model is the best of the dichotomous configurations. This index says that the mean distance between points in separate clusters is 3.5 times the mean within-cluster distance. The highest value for the other two-cluster configurations is 2.7. The good showing on this index for the collar-line model reflects primarily greater cluster compactness, with the average within-cluster mobility distance of 1.0.

There are three-cluster and four-cluster configurations with slightly higher values for the mean-distance individuation index. These configurations are all hierarchically related by subdivision to the collar line model.

Two of the three-cluster configurations have a value of 3.8 for the mean-distance index and 0.9 for the gap index. One combines the isolated top and the collar-line models (A, BC, DE). The other combines the collar-line model with a modified version of the isolated top model, which would have the top broadly defined to include both occupations A and B (AB, C, DE). A third three-cluster configuration has a considerably lower mean-distance index, 2.8, but a slightly higher gap index, 1.1. This configuration (ABC, D, E) combines the collar-line model and the isolated bottom model.

Two configurations are tied for the best four-cluster scheme with 3.9 for the mean-distance index and 1.0 for the gap index. One combines the isolated bottom model with the isolated top and the collar-line model (A, BC, D, E); the other substitutes the modified isolated top for the isolated top (AB, C, D, E).

To reinforce the conclusion that the most important division in the labor force corresponds to the distinction between white-collar and blue-collar workers, the fourteen configurations may be partitioned into a set which honors this division and a set which does not. For the seven configurations incorporating the collar-line distinction, the average value of the mean-distance index is 3.6 and always above 2.8, whereas for the other seven configurations this individuation index averages 2.2 and is never greater than 2.7. For the gap individuation index, the average value for the first set of configurations is 1.1 with a minimum of 0.9, and for the other set the index averages 0.4 with a maximum of 0.5.

In sum, there is some support from the clustering of occupations on the basis of the Laumann mobility table of an isolated top category and, to a lesser extent, an isolated bottom category of workers. Both of these features, however, are clearly secondary to the collar-line distinction.

REFERENCES

Anderberg, M. R.

1973 *Cluster Analysis for Applications* in *Probability and Mathematical Statistics*, 19. New York: Academic Press.

Arrow, K. J.

1973 "Higher education as a filter." *Journal of Public Economics*, 2 (July): 193-216.

Barber, B.

1957 *Social Stratification: A Comparative Analysis of Structure and Process.* New York: Harcourt, Brace, and World.

Becker, Gary S.

1957 *The Economics of Discrimination.* Chicago: University of Chicago Press.

Blau, Peter M. and Otis Dudley Duncan.

1965 "Some preliminary findings on social stratification in the U. S." *Acta Sociologica* 9:4-24. Reprinted in C. S. Heller (ed.), *Structured Social Inequality.* New York: Macmillan, Inc.

——— .

1967 *The American Occupational Structure.* New York: John Wiley & Sons, Inc.

Bluestone, Barry

1970 "The tripartite economy: labor markets and the working poor." *Poverty and Human Resources Abstracts* 5 (July-August): 15-35.

Bogardus, E. S.
1925 "Measuring social distances." *Journal of Applied Sociology* 9 (January-February): 216-26.
Bowles, Samuel
1972 "Schooling and inequality from generation to generation." *Journal of Political Economy* 80 (May-June): S219-52.
Bowles, Samuel, and Herbert Gintis
1972 "I. Q. in the U. S. class structure." *Social Policy* 3 (November-December).
——.
1973 "I. Q. in the U. S. class structure." *Social Policy* 3 (January-February).
——.
1975 "The problem with human capital theory–a Marxian critique." *American Economic Review Proceedings* 65 (May): 74-82.
——.
1976 *Schooling in Capitalist America.* New York: Basic Books, Inc., Publishers.
Braverman, Harry
1974 *Labor and Monopoly Capital.* New York: Monthly Review Press.
Brittain, John A.
1977 *The Inheritance of Economic Status.* Washington, D. C.: The Brookings Institution.
Cain, Glen C.
1975 "The challenge of dual and radical theories of the labor market to orthodox theory." *American Economic Review* 65: 16-22.
——.
1976 "The challenge of segmental labor market theories to orthodox theories: a survey." *Journal of Economic Literature* 14: 1215-57.
Carter, Michael A., and Martin Carnoy
1974 "Theories of labor markets and worker productivity." Discussion Paper 74-4 presented at the Center for Economic Studies, Stanford, California (August).
Centers, R.
1961 *The Psychology of Social Classes: A Study of Class Consciousness.* New York: Russell & Russell.
Council of Economic Advisors
1973 "The economic role of women." *United States President, Economic Report 1973.*
Dobb, M.
1973 *Theory of Values and Distribution Since Adam Smith: Ideology and Economic Theory.* New York: Cambridge University Press.
Doeringer, Peter B., and Michael J. Piore
1971 *Internal Labor Markets and Manpower Analysis.* Lexington, Mass.: D. C. Heath & Company.
Duncan, Otis Dudley
1961 "A socioeconomic index for all occupations." Pp. 109-38 in A. J. Reiss, Jr. (ed.), *Occupations and Social Status.* New York: Free Press.
——.
1969 "Inheritance of poverty or inheritance of race?" Pp. 85-110 in D. P. Moynihan (ed.), *On Understanding Poverty: Perspectives from the Social Sciences.* New York: Basic Books, Inc., Publishers.

Edwards, Alba M.
1934 *Comparative Occupation Statistics for the United States.* Washington, D. C.: U. S. Government Printing Office.

Edwards, Richard C., Michael Reich, and David M. Gordon (eds.)
1975 *Labor Market Segmentation.* Lexington, Mass.: D. C. Heath & Company.

Featherman, David L., and Robert M. Hauser
1978 *Opportunity and Change.* New York: Academic Press.

Feldstein, M. S.
1958 *Lowering the Permanent Rate of Unemployment.* U. S. Congress Joint Economic Committee (September 18).

Fisher, W. D.
1958 "On grouping for maximum homogeneity." *Journal of the American Statistical Association* 53:789-98.

_____.
1969 *Clustering and Aggregation in Economics.* Baltimore, Maryland: Johns Hopkins University Press.

Foley, D. K.
1975 "Problems versus conflicts: Economic theory and ideology." *American Economic Review Proceedings* 65 (May):231-36.

Freedman, Marcia
1969 *The Process of Work Establishment.* New York: Columbia University Press.

_____.
1976 *Labor Markets: Segments and Shelters.* Montclair, New Jersey: Allanheld, Osmun & Co., Publishers, Inc.

Friedman, M.
1953 "The methodology of positive economics." Pp. 3-43 in M. Friedman (ed.), *Essays in Positive Economics.* Chicago: University of Chicago Press.

Gallaway, Lowell E.
1966 "On the importance of 'Picking One's Parents.' " *Quarterly Review of Economics and Business* 6 (Summer): 7-15.

Glass, D. V.
1954 *Social Mobility in Britain.* London: Routledge & Kegan Paul.

Goldthorpe, J. H.
1964 "Social stratification in industrial society." Pp. 97-112 in P. Halmos (ed.), *The Development of Industrial Society.* New York: Humanities Press.

Goodman, L. A.
1964 "Simultaneous confidence limits for cross-product ratios in contingency tables." *Journal of the Royal Statistical Society, Series B* 26:86-102.

_____.
1965 "On the statistical analysis of mobility tables." *American Journal of Sociology* 70 (July):564-85.

_____.
1968 "The analysis of cross-classified data: Independence, quasi-independence, and interactions in contingency tables with or without missing entries." *Journal of the American Statistical Association* 63:1091-1131.

_____.
1969 "How to ransack social mobility tables and other kinds of cross-classification tables." *American Journal of Sociology* 75 (July):1-40.

Gordon, David M.
1972 *Theories of Poverty and Underemployment.* Lexington, Mass.: D. C. Heath & Company.

———.
1971 "Class, productivity, and the ghetto: a study of labor market stratification." Ph.D. dissertation. Harvard University.

Gramlich, E. M.
1974 "The distributional effects of higher unemployment." *Brookings Papers on Economic Activity* 2:293-336.

Harrison, Bennett
1972a *Education, Training, and the Urban Ghetto.* Baltimore, Maryland: Johns Hopkins University Press.

———.
1972b "Education and underemployment in the urban ghetto." *American Economic Review* 62 (December):796-812.

Heller, C. S.
1969 *Structured Social Inequality: A Reader in Social Stratification.* New York: Macmillan, Inc.

Hollingshead, A. B.
1975 *Elmstown's Youth and Elmstown Revisited.* New York: John Wiley & Sons, Inc.

Hope, K.
1972 *The Analysis of Social Mobility.* New York: Oxford University Press.

Hughes, E. C.
1959 "The study of occupations." P. 453 in R. K. Merton et al. (eds.), *Sociology Today.* New York: Basic Books, Inc., Publishers.

Jackson, M., and E. B. Jones
1973 "Unemployment and occupational wage changes in local labor markets." *Industrial and Labor Relations Review* 26 (July):1135-45.

Jardine, H., and R. Sibson
1971 *Mathematical Taxonomy.* New York: Wiley-Interscience.

Jencks, Christopher S., M. Smith, J. Acland, M. J. Bane, D. Cohen, H. Gintis, B. Heyns, and S. Michelson.
1972 *Inequality: A Reassessment of the Effect of Family and Schooling in America.* New York: Basic Books, Inc., Publishers.

Johnson, S. C.
1967 "Hierarchical clustering schemes." *Psychometrica* 32:3 (September): 241-54.

Kalecki, Michael
1970 *Selected Essays on the Dynamics of the Capitalist Economy, 1933–1970.* New York: Cambridge University Press.

Keat, P. G.
1960 "Long-run changes in occupational wage structure, 1900–1956." *Journal of Political Economy* 68:584-600.

Kerr, C., et al.
1960 *Industrialism and Industrial Man: The Problems of Labor and Management.* Cambridge, Mass.: Harvard University Press.

Kohn, M. L.
1969 *Class and Conformity: A Study of Values.* Homewood, Illinois: Dorsey Press.

Kuhn, T. S.
1970 *The Structure of Scientific Revolutions.* Chicago: University of Chicago Press.

Kullbach, S.
1968 *Information Theory and Statistics.* New York: John Wiley & Sons, Inc.

Landecker, W. S.
1960 "Class boundaries." *American Sociological Review* 25 (December):868-77.

Laumann, E. O.
1966 *Prestige and Association in an Urban Community: An Analysis of an Urban Stratification System.* Indianapolis: The Bobbs-Merrill Co., Inc.

Lipset, S. M., and R. Bendix
1959 *Social Mobility in Industrial Society.* Berkeley: University of California Press.

Maher, J. E.
1961 "The wage pattern in the U. S., 1956–1957." *Industrial and Labor Relations Review* 15:3-20.

Mannheim, K.
1936 *Ideology and Utopia: An Introduction to the Sociology of Knowledge.* New York: Harcourt, Brace & Co.

Marglin, S. A.
1975 "What do bosses do? Part II." *Review of Radical Political Economics* 7 (Spring):20-37.

Marshall, A.
1947 *Principles of Economics,* 8th edition. New York: The Macmillan Company.

Marx, Karl
1962 *Capital: A Critique of Political Economy.* Moscow: Foreign Languages Publishing House.

Matza, D.
1966 "The disreputable poor." Pp. 310-39 in N. J. Smelser and S. M. Lipset (eds.), *Social Structure and Social Mobility in Economic Growth.* Chicago: Aldine Publishing Co.

Mincer, J.
1970 "The distribution of labor incomes: a survey with special reference to the human capital approach." *Journal of Economic Literature* 8 (March):1-26.

Mills, C. Wright
1959 *The Sociological Imagination.* New York: Oxford University Press.

Montagna, Paul D.
1977 *Occupations and Society: Toward a Sociology of the Labor Market.* New York: John Wiley & Sons, Inc.

O'Connor, James
1973 *The Fiscal Crisis of the State.* New York: St. Martin's Press.

Okun, A. M.
1973 "Upward mobility in a high-pressure economy." *Brookings Papers on Economic Activity* 1:207-52.

Ossowski, S.
1963 *Class Structure in the Social Consciousness.* Translated by Sheila Patterson. London: Routledge & Kegan Paul.

Piore, Michael
1969 "On-the-job training in the dual labor market." Pp. 101-32 in A. R. Weber, F. Cassell, and W. L. Ginsberg (eds.), *Public-Private Manpower Policies.* Madison: Industrial Relations Research Association, University of Wisconsin.

——.
1970 "Jobs and training." Pp. 53-83 in S. H. Beer and R. E. Barringer, (eds.), *The State and the Poor.* Cambridge, Mass.: Winthrop Publishers, Inc.

——.
1975 "Notes for a theory of labor market stratification." Pp. 125-50 in R. Edwards et al., *Labor Market Segmentation.* Lexington, Mass.: D. C. Heath & Company.

Rainwater, L.
1968 "The problem of lower-class culture and poverty-war strategy." Pp. 229-59 in D. P. Moynihan (ed.), *On Understanding Poverty: Perspectives from the Social Sciences.* New York: Basic Books, Inc., Publishers.

Reder, M. W.
1955 "The theory of occupational wage differentials." *American Economic Review* 45:833-52.

Reiss, A. J., et al.
1961 *Occupations and Social Status.* New York: The Free Press.

Roberts, N. J.
1974 "On the nature and condition of social science." *Daedalus* 103 (Summer): 47-64.

Rogoff, N.
1953 *Recent Trends in Occupational Mobility.* New York: The Free Press.

Rosenberg, Samuel
1975 "The dual labor market." Ph.D. dissertation. University of California at Berkeley.

Schultz, Theodore W.
1961 "Investment in human capital." *American Economic Review* 51 (March): 1-17.

Schumpeter, J. A.
1954 *The History of Economic Analysis.* New York: Oxford University Press.

Smith, R. E., J. E. Vanski, and Charles C. Holt
1974 "Recession and the employment of demographic groups." *Brookings Papers on Economic Activity* 3:737-60.

Stiglitz, J. E.
1973 "Approaches to the economics of discrimination." *American Economic Review Proceedings* 63 (May):287-95.

——.
1974 "Equilibrium wage distributions." *Technical Report 154.* Stanford: Institute for Mathematical Studies in the Social Sciences.

Taussig, F. W.
1921 *Principles of Economics*, 3d ed. New York: Macmillan, Inc.

Thiel, Henri
1971 *Principles of Econometrics.* New York: John Wiley & Sons, Inc.

U. S. Bureau of the Census
1964 "Lifetime occupational mobility of adult males, March, 1962." *Current Population Reports,* Series P-23, No. 11 (May 12). Washington, D. C.: U. S. Government Printing Office.

Wachter, M. L.
1970 "Cyclical variation in the interindustry wage structure." *American Economic Review* 60 (March):75-84.

_____.
1975 "Primary and secondary labor markets: a critique of the dual approach." *Brookings Papers on Economic Activity* 3:637-80.

Ward, J. H., Jr.
1963 "Hierarchical grouping to optimize an objective function." *Journal of the American Statistical Association* 58:236-44.

Weber, Max
1946 *From Max Weber: Essays in Sociology.* C. Wright Mills, ed. and trans. New York: Oxford University Press.

White, H. C.
1963 "Cause and effect in social mobility tables." *Behavioral Sciences* 8:14-27.

Wright, E. O., and L. Perrone
1977 "Marxist class categories and income inequality." *American Sociological Review* 42 (February):32-55.

Zelan, J.
1964 "Recruitment to the legal profession." Ph.D. dissertation. University of Chicago.

_____.
1972 "Social origins and the recruitment of American lawyers." In R. M. Pavalko (ed.), *Sociological Perspectives on Occupations.* Itasca, Illinois: F. E. Peacock Publishers, Inc.

INDEX

University Presses of Florida
Gainesville

0-8130-0674-0