WORKING PAPER

MARKETS AS QUEUES, WITH AN APPLICATION TO EDUCATION

Ross Boylan

February 1989
WP-89-018

IIASA
International Institute for Applied Systems Analysis

MARKETS AS QUEUES, WITH AN APPLICATION TO EDUCATION

Ross Boylan

February 1989
WP-89-018

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

Foreword

An important interface between economics and demography has been the issue of the effect of size on the relative welfare of one group vis-a-vis another. Well-known examples are the Easterlin framework for fertility dynamics and the work of Freeman, Welch, and others on the impact of birth cohort size on earnings differentials. The basic hypothesis underlying their work is that being a member of a larger birth cohort can have a deleterious effect on one's life prospects, at least when these are expressed in terms of income or consumption aspirations. In this paper, Ross Boylan uses a queuing theoretic approach to examine the relationship between income and the size of groups defined in terms of educational attainment, and more specifically, whether an individual holds a particular credential such as a high school diploma or a college degree. In contrast to the work cited above, Boylan hypothesizes that, if markets work like queues, with an implicit matching between particular types of jobs and individuals with credentials, then an increase in relative group size may result in an improvement in the relative welfare of the group.

Boylan provides the following example: Suppose there are two kinds of jobs, middle and low income (say, \$10 and \$3 per hour), and that people either have a credential or not. Initially, credential holders all have \$10 jobs, while the uncredentialled are split between low and middle income jobs. Then more people get credentials. As a result, some of the uncredentialled are driven out of the \$10 jobs, but all those with credentials remain middle income. Thus, the income of those with credentials remains unchanged, but the average income of those without credentials falls. In the paper, Boylan generalizes from this simple model to that of a continuous distribution of income and more broadly defined credential groups.

Empirically, the effects of credential group size on relative earnings are quite small. Nevertheless, one can imagine that with further development the queuing model might provide some additional insight about the earnings differentials that are often observed to exist between groups defined in terms of other demographic characteristics such as sex or race.

Charles A. Calhoun
Research Scholar and
Acting Deputy Program Leader
Population Program

Contents

Markets as Queues, with an Application to Education

Ross D. Boylan*

Abstract

Sociologists often claim that larger structures, such as labor markets, can not be understood simply as an aggregation of individual exchanges. Yet most models we use are individualistic. This paper develops a queuing model which separates the distribution of individual characteristics from the structure of jobs, allowing a role for each. The model thus makes a sharp distinction between the value of a characteristic for an individual and its value for society, and eliminates the need to assume that people are paid what they're worth. In such a model a group may *improve* its standing relative to others as its size increases. This possibility is investigated as an explanation for the disproportionate income gains of those with high school and college diplomas. Analysis of some recent U.S. data with a queuing model show small effects of group size on these income gains, although regression based analysis might lead one to conclude that size effects are large.

*I would like to thank James Coleman, Nathan Keyfitz, James March, Robert Mare, John Meyer, and Nancy Brandon Tuma, for their comments on the work here. Many other participants at colloquia also gave me valuable feedback. The International Institute for Applied Systems Analysis provided facilities to complete this work. The Inter-University Consortium for Political and Social Research provided this study's data, which originated with the U. S. Department of Commerce, Bureau of the Census. Stanford University provided computer resources. None of them bears any responsibility for the analyses or interpretations herein.

This paper focuses on a narrow question, which grows out of some more general concerns. The narrow question is whether the size of different educational groups affects the income of those groups. This question stems from concerns about how education operates in the labor market and, indeed, how the labor market operates in general. And behind these concerns lie general issues of modeling social structure, combining micro and macro processes, and the appropriateness of regressions as theory-testing tools. This paper's general theme is that a familiar sociological truth—that social structure counts—has yet to penetrate many of our mathematical models or empirical analyses very far, and that a rigorous working through of this truth will yield some surprises and some benefits. The particular approach taken below intends to illustrate that point, rather than to provide the definitive answer to the question of how the labor market works.

People with more education make more money (Blau and Duncan, 1967; Featherman and Hauser, 1978; Jencks and *et al.*, 1979; Kelley, 1973; Mincer, 1974), not only in the U.S. but worldwide (Heyneman and Loxley, 1983). Although neoclassical economics takes this as a sign that the more educated are more productive, the evidence suggests productivity differences caused by education, if they exist at all, are much less than pay differences. This conclusion follows both from individual level studies and aggregate dynamic analyses (Berg, 1970; Berg et al., 1981; Collins, 1971; Collins, 1979; Lockheed et al., 1980; Benavot, 1985). These results fit the general finding that productivity, whatever its causes, seldom matches income (Frank, 1984; Medoff and Abraham, 1981; Thurow, 1968). Of course, measures of productivity apart from pay are difficult to obtain, and all studies which attempt such measurement are open to criticism. But this measurement difficulty suggests an important theoretical problem, for the best developed theory of income determination, human capital theory of neoclassical economics, holds that productivity determines income. The scientific status of the theory is thus questionable, since it rests so heavily on an unobservable. Further, something which agents generally can not observe will only determine pay if some strong assumptions are true. In contrast, the model below requires only that employers know which kind of people they would prefer to hire (given a job with a fixed wage rate), and that employees know which kind of jobs they would prefer to take.

Discussion of race and sex discrimination often treats these as aberrations in an otherwise meritocratic market which pays people what they are worth. However, the results above suggest the basic view of how markets function needs to be reconsidered, and considerable evidence supports this claim (Albelda et al., 1987; Granovetter, 1985; Kalleberg and Sørensen, 1979; Lester, 1954; Piore, 1979; Thurow, 1975; Thurow, 1983). Economists are aware that the neoclassical theory of the labor market has problems, and have suggested a variety of ingenious repairs.[1] However, all hold that in some average sense groups receive the value of what they produce. Given that this assumption appears to contradict the scattered available evidence, it seems worthwhile to pursue models which do not make such an assumption.[2]

How might markets work, if not by rewarding productivity? One possibility is that characteristics such as race, gender, age, and education have a social value, determined by political struggles and other institutional processes. In this view, social structure enters into the market by setting the rules under which the market operates, for example, "pay $1,000 more for a high school diploma." However, given a set of rules, the total market outcome is simply the aggregation of the individual outcomes; if the people who were on the market changed (for example, if more college graduates entered the market) the aggregate outcome would change (more people would earn more money).

Another possibility is that a set of positions exist, and that people and

[1] Screening theory suggests that employers use education and other traits as signals of productive ability, even if the signals do not cause productive ability (Spence, 1974). More recent efficiency wage arguments propose several reasons that productivity might depend on pay (e.g., higher wages reduce turnover, aid discipline, boost morale, satisfy equity considerations, impede unionization, promote health, save on contract renegotiation expense) (Bulow and Summers, 1986; Stiglitz, 1984; Yellen, 1984).

[2] See the above literature for some of the arguments on this point. It may be useful to highlight some of the theoretical reasons competitive market pressures might not equate pay and marginal product. First, the limits of human and organizational rationality imply that competition will never be "perfect," and that untrue beliefs may become widely shared. Second, various institutional processes shape the market quite apart from the actions of supply and demand. Third, productivity may be so intricately bound to the social world (so that only certain kinds of people are deemed appropriate for certain roles) that the technical concept of a marginal product of an asocial producer can be seriously misleading.

groups compete for these positions. Social structure enters these models in a double sense: first, by determining the rules of the competition, and, second, by determining the positions which exist. From this perspective, the aggregate market is *not* the sum of the individual exchanges in it—if the individuals on the market changed, the total distribution of jobs would remain the same. More modestly, the processes which distribute individuals among jobs are not tightly linked to those which determine the aggregate structure of jobs. This claim is the thrust of much recent structural theorizing about the labor market, and this paper adopts that perspective. However, the focus of this paper differs from that of much structural research, and those differences and their significance merit some discussion. With that background, the paper proceeds to develop a simple structural model of the the labor market, and uses that model to analyze the relation between education, group size, and income. These results are contrasted with those from regression analysis, and a conclusion returns to some of the larger issues behind the analysis.

1 The Need for a Structural Account of Individual Success

This paper focuses on individual attainment, and thus differs from much recent structural research. Indeed, the "new structuralism" was born in an explicit effort to shift the emphasis of stratification research from the study of individual success and failure to the study of the structures or "empty spaces" occupied by people. Different analysts emphasize different forces shaping these structures—political, organizational, institutional, or economic—but all agree that the structures can not be understood simply as aggregations of individual exchanges.[3] This general claim is an old one is sociology.

However, the structural account is incomplete without an account of individual attainment, for the attainments of individuals, taken collectively, must match the structure of the market even if the market can not meaningfully be understood as arising from those attainments. There lies the problem which motivated the development of the model below: existing models of individual attainment provide one account of what the aggregate market looks like

[3]See Baron, 1984 and Althauser, 1987 for recent reviews.

(the whole is simply the sum of the individual outcomes); structural theories provide another (the whole is the product a various larger political, organizational, institutional and economic forces); and the two accounts will not generally match.[4] The structural account of the labor market will not be complete until it can offer an alternative account of how individuals are distributed among positions in the economy.

These claims may seem surprising, for structural studies *have* examined individual attainment, even if this has not been their main concern. However, the methods used have generally been variants on individualist ones. The typical structural study of individual attainment shows that the relationship between personal characteristics and outcomes depends on structural position, for example that education is rewarded more strongly in some sectors than others (Beck et al., 1978; Stolzenberg, 1978; Dickens and Lang, 1984, are representative). As demonstrations that structure counts, these exercises are quite successful; as complete models of attainment, they are less satisfactory, for two reasons. First, the distribution of people across sectors is often assumed rather than modeled, so there is no way of saying how the distribution would change if the people or the structure changed. Second, outcomes which are modeled, whether sectoral location, income, or employment, are modeled in a way which undercuts the structural perspective. That is, the outcome is a function of the characteristics of the person, as in a regression equation. This specification implies that if the aggregate characteristics of the people on the market changed, the aggregate outcomes—the structure of the market—would change in an exactly corresponding fashion. So such models are less structural than they appear.

Of course, many of the investigators probably did not intend that their analyses be taken quite so rigorously as models. They wished to show that structure mattered, and they did so; they did not intend that the estimates be taken as parameters of some underlying process which would remain constant while the structure or demography of the market changed. On the other hand, some may have intended a sort of soft structural model, one in which there

[4]Structural models of individual attainment do exist, most notably in models of vacancy chains (Sørensen, 1977; White, 1970). Also, models of mobility tables generally assume that people are distributed across a fixed set of positions (Boudon, 1974; Featherman and Hauser, 1978). However, none of these models has the generality of regressions, which can describe how an arbitrary set of characteristics affect a continuously distributed outcome.

was no rigidly pre-existing structure of jobs and income. The models above may be appropriate for that.

However, the central theme of much of the verbal theorizing about the labor market is that there is a fairly sharp separation between the processes generating the job structure and those which distribute people among jobs. It is common to think of the labor market this way: for example, a changing political climate and the internationalization of capital has eliminated many middle-income jobs; the baby boom has entered the market. Yet we do not have a very rigorous way of tracing through the distributional effects of these changes on different groups; the models most commonly used undercut the structural perspective which motivates the questions. Thus a truly structural account of individual attainment is needed for more than logical completeness of structural theory; it is needed to understand and predict how different groups will fare under changing conditions. The following sections illustrate one such model and contrast it with more conventional regression models.[5]

2 A Queuing Model

2.1 Background

This section develops a model which formalizes some existing theories of the labor market. Structural analyses of the labor market hold that jobs exist apart from the people who fill them, and that the market is an arena in which jobs and people match (Granovetter, 1981). Many models would be consistent with that view; this paper pursues a particularly fruitful and simple one. Thurow (1975) suggested regarding the labor market as two queues: one queue has jobs, ranked by their desirability to employees; the other queue has people, ranked by their desirability to employers.[6] Importantly, wages attach to jobs rather than people. The model below makes the additional assumption that jobs are ranked by their wages. An error term will account for the approximation involved in this assumption. These assumptions are undoubtedly heroic, more

[5]The general neoclassical model does have a supply and demand side. While such an approach can at least partially incorporate structural and demographic change, it is not the only way to do so. For reasons given above, this paper pursues an alternative account.

[6]Thurow emphasized training costs for ranking people; the approach here is more agnostic.

applicable to some parts of the market than others. A full defense of their plausibility would require a recapitulation of the debates cited above; suffice it to say these assumptions seem as reasonable a way of characterizing the labor market those of the neoclassical model. The intent of this article is not to prove one model right or wrong, but to show that the question of how the market works is critical, that exact specification of an answer to that question has some benefits, and that more structural models of the market can be developed.

This verbal formulation of a queuing model suffices to make some points. The image of a queue suggests that if one person moves up in the queue, for example by getting more education, he or she does so by bumping other people down in the queue. The assumption that wages attach to jobs, which exist apart from people, implies that people do not necessarily get what they're "worth." Like all structural approaches, this one suggests viewing wealth and poverty as the result of social processes rather than as the result of individual virtue or deficiency. This viewpoint has immediate implications for virtually all labor market policy research: inferences from the micro to the macro level may not be warranted. For example, a job training program raises the income and employment of participants. Would it affect poverty if widely applied? Not if participants are simply displacing others into less desirable positions. Another example: a welfare program reduces labor force participation. If widely applied, would it reduce economic activity? Not if participants are simply yielding positions to others. The widespread practice of applying micro relationships to yield macro estimates of the impact of government programs rests on unexamined assumptions about the market.

However, words alone can not answer some other questions. The verbal model makes clear that the distribution of income may change apart from any changes in individuals (e.g., a recession, or deindustrialization, or an assault on unions), but it doesn't predict how different groups will fare. Nor does it help in predicting the effects of population change (e.g., a baby boom or a rise in the number of college graduates). Nor does it predict the effects of changes in the matching and ranking process, such as a decline in discrimination. For that a formal mathematical model is required, and to that we now turn.

2.2 The Formal Model

The model here requires three things:

- a set of jobs, with wages
- a set of people, with characteristics
- a rule for ranking people.

Given these three things, the model matches the people to the jobs. Since the rule for ranking people has a random element, the model can not predict precisely where an individual or group will end up, but (like all probabilistic models) it predicts the likely distribution of outcomes.

The operation of the model is simple: jobs are ranked by their wages; people are ranked according to the ranking procedure (details below); and the people and jobs are matched. The top person gets the top job; the second-ranked person gets the second-ranked job, and so on. If there are more people than jobs, the lowest ranked people are unemployed.

The last detail is the ranking rule, a device to express the fact that personal characteristics influence queue position, but they do so somewhat randomly. The rule specifies that ranking proceeds as follows:

1. Assign everyone a score
2. Rank everyone by the scores they receive.

Scores are random functions of personal characteristics; mathematically

$$s_i = f(\mathbf{X}_i) + \epsilon_i, \tag{1}$$

where s_i is the score for individual i, $\mathbf{X}_i$ is a vector of characteristics (e.g., age, education, race, sex), ϵ_i is an error term, and f is a function relating the characteristics to the score. The function and the error term must be specified (or estimated) before operation of the model. Equation (1) will be called the *score equation.*

2.3 Interpretation

The score is an index of overall desirability of a person to employers. Clearly, the model will be more successful when employers have a consensual ranking of individuals, although certain departures from consensus are captured by the error term. However, it is important to emphasize that the score most likely is *not* a direct measure of quality or productivity; it can just as well reflect prejudice or unfounded beliefs about the value of certain traits.

The error term reflects unobserved differences among individuals, but it also reflects chance: identical people may get different jobs simply because they arrive on the market at different places or times.

3 An Application: Education, Credentials, and Group Size

This section presents an application of the queuing model developed above. The goal is both to show that the model can fit the data, and to illustrate the new kinds of theoretical issues which the model raises. Unlike neoclassical economics, the queuing model suggests that larger groups may have relatively better outcomes in the labor market, and this paper investigates the relationship between the exceptional size of certain educational groups and their exceptional income.

3.1 Theoretical Discussion

Analysis focuses on a particular puzzle, credentials effects, which analysts have thought might illuminate some more general features of the labor market. Credentials effects are the additional gains in income and status associated with a holding a certificate of graduation. Most analysts find that the year of school in which one gets a certificate is worth more than other years of schooling (Blaug, 1976; Jencks and *et al.*, 1979; Layard and Psacharopoulos, 1974). This finding has provoked considerable controversy, for it suggests that employers reward certificates rather than underlying productive ability (which, it seems reasonable to assume, schools impart in a fairly continuous manner). All sides to the debate agree that if certificate holders get more it is because there is something distinctive about them or the certificates they hold.

The debate over credentials has focused only on the distinctiveness of the credentials or the credential holders, but has overlooked the demography of the market. Consider Figures 1 or 2. These describe the two samples analyzed below, but typify the whole population. The solid line shows the average earnings of each educational group; it takes sharp jumps at high school and college completion.[7] The dashed histogram shows the number of people at each grade level; it too jumps at the credential points, particularly 12'th grade, high school diploma. Could the large size of these groups be a cause of the income jumps?[8]

Indeed, general theoretical debate over education has emphasized the sheer role of numbers. Some have argued that the value of education is relative, rather than absolute. A high school diploma may put one at the top of the labor market if no one else has such a diploma, or at the bottom if everyone else has gone to college. The value of any particular level of education depends on the size distribution of the different educational groups. This is precisely the viewpoint of the queuing model: the value of a given score depends on the distribution of the scores of others.

Thus, this paper will attend particularly to group size. Note that the population in question is those in the labor market; the groups consist of those with different levels of education. *Throughout this paper group size means the number of people on the market who have completed a given level of schooling. Size effects refer to the effect of changing group size on the income of that group.*

Neoclassical economics suggests that larger groups should be at a disadvantage on the labor market. The failure of returns to college to fall as the number of college graduates increased has been emphasized by critics of the neoclassical theory (Boudon, 1974; Thurow, 1975), and a queuing perspective can give a radically different analysis of group size.

How? Suppose that employers place *no* special value on diplomas, but they value each year of education in a strong sense: they always prefer to hire someone with more education to hiring someone with less. Thus all those with n years of education will rank above all those with less education in the queue of people.[9] The ranking this produces is shown in Figure 3, with people

[7] As subsequent analysis and the studies cited above show, this pattern persists after controls.

[8] It is plausible that people are drawn to these schooling levels by the associated benefits; it is also plausible that widely held beliefs that it is good to graduate cause both the concentration of people and the jump in income. The analysis below asks what the consequences of size are, regardless of the source of that size.

[9] This assumes all people are on the market at once and employers have perfect information. These assumptions serve only to simplify exposition, and will be dropped in the empirical analysis.

ranked in a queue on the right and jobs ranked in a queue on the left. For simplicity, assume the distribution of income is uniform. In that case, the mean income of each educational group is shown by the horizontal bars in the center of the figure. The vertical distance between these bars indicates the mean income gain associated with completing each additional year of school. The crucial point is that the gain associated with going from grade 11 to 12 is greater than the gain from 10 to 11. The larger gain is the result of the large number of twelfth graders, not the special treatment by employers of high school graduates. 12'th graders are an unusually large group, and so occupy an unusually large range of jobs and income. Thus the jump in average income between 11'th and 12'th grade is exceptional.

The queuing framework thus introduces a distinction between the effects of characteristics on the matching process and their effects on outcomes. In this case, credentials may have no special effect in matching while retaining a large effect on outcomes. But the effects are size effects.

Ranking effects will refer to the effect of a characteristic in the matching competition; mathematically they refer to the the relation between characteristics, such as education, and scores in the score equation (1). *Outcome effects* will refer to the association between characteristics and income.

Of course, it is quite possible that credentials do have a special effect in the matching process, so the queuing framework does not require that size effects cause the credential outcomes. It does, however, call into question the interpretation of the observed credentials effects in regressions.

The value of an additional year of school thus depends on three things:

- the distribution of job opportunities,
- the distribution of competitors and their attributes,
- the extent to which that year of schooling helps in the struggle for queue position.

Discussion so far has ignored job opportunities, but these clearly matter as well. The income gained by moving past a given number of people depends on the jobs one is moving past as well; moving up a few jobs at the top of the income distribution (which has relatively few jobs in each income interval) will produce larger income gains than moving past the same number of jobs in the middle of the income distribution.

In short, this example shows that a trait may be valuable not because of intrinsic qualities of the trait, but because of the size of the group with the trait. Since the model above, in general, traces outcomes to the interaction of jobs, demography, and the competitive value of characteristics, the relative weights of the three factors is an empirical question. Discussion now turns to the empirical analysis.

3.2 Data

Samples

This study uses two subsamples of the Current Population Survey, a household-based survey of the civilian noninstitutionalized population of the U.S.[10] The first subsample contains only those who moved between employers; the second contains movers and stayers aged 40–55; both are restricted to white males with positive earnings in 1979. Earnings data were collected for all those aged 15 and over. The March, 1980 survey (U.S. Dept. of Commerce, Bureau of the Census, 1984) provided information about labor force experience over the entire calendar year 1979, a business cycle peak with an unemployment rate of 5.8%.

The sample restrictions by race, sex, and earnings, though fairly conventional, merit brief discussion.

Race and Sex Many argue that markets are segmented by race, or sex, or both, and previous work with this model suggested such segmentation (Boylan, 1986). Analysis of separate race-sex groups seems the safest course, since throwing non-competing groups together in one market is clearly wrong, while separating competing groups may still yield valid estimates of the effects of education within groups. Small sample size precluded the analysis of blacks. Since the analysis of women would

[10]The institutionalized population generally does not compete on the labor market, so its omission should cause no problem. The military does compete with civilian jobs, and particularly near—but not at—the bottom of the labor market this is significant (Mare and Winship, 1984). It also screens on somewhat different characteristics than those used for civilian jobs. The net effect is difficult to gauge; it is likely to be largest for blacks, who are excluded from this analysis, and for youth, who are present in only one of the analyzed subsamples. So, while the omission is unfortunate, meaningful analysis of the rest of the labor force should be possible.

raise complex issues of labor force participation, this study analyzes only males.

Earnings The queuing model does not require exclusive focus on those with positive earnings, and such a restriction obviously introduces sample selection bias. Why make the restriction? First, regression models typically do employ such a restriction, and its use enhances comparability of results. Second, the restriction to those with positive earnings simplifies both the mechanics and exposition of the model.

Concern with selecting a theoretically appropriate sample drives the definition of the mover sample. In the queuing model everyone competes with everyone else for the available jobs. Many jobs inside organizations are only open to those already in the organization, and so should be excluded from the scope of the model. The mover sample meets this theoretical concern as much as is practical, by including only those who made or attempted a move between employers. However, this restriction in turn raises problems, for it reduces the sample to slightly more than 1/6 of those in the labor force, and includes many teenagers with questionable labor force attachment. As a check, parallel analyses were carried out on a mature (ages 40 to 55) male sample which excludes teenagers and includes all those in the labor force, people who stay in the same firm as well as movers. Further, to the extent that age segments markets, narrower age ranges are more appropriate for analysis. Ages 40 to 55 are the peak years of average earnings (Mincer, 1974, chapter 4). Thus, the analysis of this age group can be interpreted as covering the long run or maximal effect of education on earnings.

Table 1 reports the sizes of the different samples.

Variables

Table 2 compares the different samples across selected variables. The labor force status variables distinguish labor force participants from non-participants and movers from non-movers. The demographic information is conventional, except for experience, which is years since graduation.[11]

[11]Experience was constructed as age−(education+6), or age−14 if education was less than 8 years. In a few cases this procedure produced negative experience; it was recoded to 0 as reported in Table 3.

The natural logarithm of hourly earnings is the study's dependent variable. Hourly earnings are total earned income divided by an estimate of total hours worked (= typical hours per week times total weeks worked in 1979). The phrase "hourly earnings," rather than "wage," underlines this variable's inclusion of salary as well as wages, and its constructed nature. A small fraction of extreme values were recoded, as reported in Table 3.

Movers are younger, poorer, and work less than white males as a group, while the reverse holds for mature white males. Their educational levels are roughly equal.

3.3 Methods

Overview

I will first fit a queuing model to the data, to provide a baseline for further analysis and to demonstrate that such a model can yield a good fit. This baseline model uses the observed income distribution and observed group sizes, and it allows special effects of credentials directly in the ranking process. The remainder of the analysis separates out the effects of size from the effects of credentialling.

Recall that in the preceding example each year of education had the same ranking effect (effect on scores which determine ranking), yet each grade had different outcome effects (effects on income), so that the observed credential outcome effects were really size effects, attributable to the large number of people with credentials. That example was simplified, and really only demonstrated the theoretical possibility that size effects might matter. Empirically, how big are they?

The first variant model, after the baseline model, tests the importance of pure size effects. It does so by constraining the ranking effect of each year of education to be the same. For these data, the credentials (outcome) effects produced by this pure size model essentially do not exist. The reasons for this are discussed.

A second variant model considers the possibility that size matters in interaction with credential ranking effects. It does so by simulating a matching process in which each grade level has the same number of people. This modestly reduces the credential outcome effects. Again, size effects are weak.

Credential outcome effects stem largely from credential ranking effects.

The model is unconventional, but the analysis follows a conventional pattern: fit parameters under various specifications and then interpret them. The model is more complex than a typical regression model, so interpretation requires simulation rather than simply reading coefficients. Some people familiar with the results have expressed frustration that this is not a test of the model (see Boylan, 1986 for such an effort), but that is not the intent of this paper, which seeks, rather, to show how such a model can be used to answer the question "how big are size effects?" Like all analysis, this one assumes a certain framework, and then answers questions within it. The next major section of the paper will compare these results with those obtained from a more conventional regression framework. This section, however, concerns the mechanics and results of the queuing analysis.

Queuing Models

Analysis using the queuing model developed above poses two distinct problems: parameter estimation and parameter interpretation. Discussion takes up each in turn.

Parameter estimation for the score equation (1) faces an immediate problem: the scores in the ranking process are not observable. The solution lies in noting that the scores are an artifice for saying that one group or individual ranks above or below another. Any set of scores which retain the same ordering of people will do. By assumption, the wages of a job reflect its ranking and so that of the person who occupies the job. So it is sufficient to estimate a regression of log wages on personal characteristics, and then reinterpret the dependent variable as a score. Thus, the queuing model interprets regression parameters as measuring relative standing rather than pay, and the relation between traits and scores, rather than between traits and income, is assumed constant. To estimate the parameters of a score equation it suffices to estimate the corresponding regression of wages on personal characteristics.[12]

Parameter interpretation proceeds by simulation. This simulation procedure is essential, since the parameters of the model yield only the effects of characteristics on ranking, not on outcomes (income). Given parameter estimates and a set of people and jobs, the model can be simulated on the

[12]See the appendix of Boylan, 1986, for a rigorous justification of this procedure.

computer to yield a possible matching of people and jobs. Since the simulation has a random element, multiple simulations are needed to uncover the average tendency. Each simulation yields a matching of people to jobs, and a conventional regression is performed on this simulated sample to summarize the relation between personal characteristics and income. This analysis will focus on the education-income relationship, particularly on the size of credential dummy variables. The mean and standard deviation of each coefficient across all simulations will be reported.

Readers without a taste for technical detail may wish to skip to the next subsection, or even to the results. A more detailed description follows.

At this point it will be useful to introduce some notation and some regression models which will figure in the analysis. The regression models have two roles: first, they will summarize both the actual data and the results of model simulation; second, some of them will form the basis for the score equation which describes the ranking process in the model.

Let S refer to years of schooling, X to years of experience, and D_n to a dummy equal to 1 for those with n or more years of schooling, 0 otherwise. $\log(w)$ is the natural logarithm of hourly earnings (w for wage). A typical regression model is

$$\log(w) = \alpha_0 + \alpha_1 S + \alpha_2 S^2 + \alpha_3 X + \alpha_4 X^2 + \alpha_5 SX + \epsilon \qquad (2)$$

S^2 and X^2 are squares of the corresponding terms; SX is schooling times experience; ϵ is an error term.

Equation (2) is one of the more flexible forms used in one of the major empirical works in the human capital tradition (Mincer, 1974, p. 92). If that equation is further simplified, so that $\log(w)$ is a linear function of education alone, it becomes "the fundamental empirical relationship" used in the human capital literature (Rosen, 1977, p. 12).

Credentialists add dummy variables for diplomas:

$$\begin{aligned} \log(w) &= \alpha_0 + \alpha_1 S + \alpha_2 S^2 + \alpha_3 X + \alpha_4 X^2 + \alpha_5 SX + \\ &\quad \gamma_1 D_8 + \gamma_2 D_{12} + \gamma_3 D_{16} + \epsilon \qquad (3) \end{aligned}$$

γ_2, for example, could be called a "high school diploma effect." If the γ are statistically significant, then there is something to be explained. While it is common to interpret them directly as credential effects, the arguments above suggest some portion of these effects may be size effects. A regression of this form will be used to summarize the relationship between education and income for both the observed and simulated samples.

A typical queuing analysis proceeds as follows: first, one specifies a score equation. Analysis below employs two equations, one like equation (3) with score as the dependent variable, and the other

$$s = \alpha_0 + \alpha_1 S + \alpha_3 X + \alpha_4 X^2 + \epsilon. \tag{4}$$

The former equation allows the ranking effects of each year of education to vary with grade level, and includes credential ranking effects. The latter equation holds that s, the score used in ranking, is affected equally by each year of schooling ($\alpha_1 S$). This differs from the other specification in omitting quadratic and interaction terms involving schooling.

With this background, a detailed description of a prototypical parameter estimation and interpretation is possible.

Estimate the parameters of the score equation by estimating a regression with log wages as the dependent variable. Simulation can then proceed, given a set of people and jobs. Generally the observed sample is used, but for one analysis, described below, this is not the case. In any case, the simulation ignores information about which people have which jobs and provides its own matching of people and jobs.

Each simulation proceeds as follows:

1. Rank the jobs by their wages, with the highest paying ranked first.
2. Rank the people.
   (a) Assign each person a score according to the score equation and parameters. This equation includes an error term; draw a random variable with the distribution of the error term (by computer simulation) and add this to the rest of the equation.
   (b) Once all people have scores, rank people by their scores.
3. Match the people to jobs, top-ranked person to top-ranked job, second person to second job, and so on.[13]

[13]Usually some people would not have jobs, but these simulations concern samples of employed people only.

4. Estimate the coefficients of equation (3) for this simulated sample and record the coefficients. D_{12} and D_{16}, the credential dummies, will be of most interest.

All analyses employ 200 simulations. The results consist of the mean and standard deviation across all simulations of each coefficient (e.g., D_{12}) estimated in the last step.[14]

Alternative Models

The "basic simulation" uses equation (3), which has credential dummies, as the score equation and uses the observed sample of people and jobs as the basis for simulation.

The "No Credentials" model uses equation (4), in which each year of education has the same ranking effects, as the score equation. Simulation employs the observed sample of people and jobs. Any credentials effects observed in such a model stem purely from the interaction of group size and the income distribution.

The "Equal Sizes" model uses the same score equation as the basic simulation, but employs simulations in which each group has the same size.[15]

The models explore the importance of group size in producing the observed credentials effects. The No Credentials model asks how important size effects are without any credential ranking effects, while the Equal Size model asks how large size effects are in the presence of credential ranking effects.

3.4 Results

Tables 4 and 5 report the results of the analysis for the samples of movers and mature males respectively. The first column, labelled "Actual Data," reports the observed relation between education, experience, and income in

[14]The standard deviation simply shows how precise these estimates of mean values are. They refer to uncertainty concerning the simulations, not uncertainty stemming from the fact that we have one of many possible samples.

[15]Actually, only the population above grade 6 was equalized, since there were so few cases below that level. For each simulation a hypothetical sample of people was drawn. This sample had exactly the same people as in the original sample below grade 7. Above grade 6, the original sample in each grade was sampled with replacement n times, where n is 1/12 the total observed sample size in grades 7–18.

the samples. It is the credential dummies, D_{12} and D_{16} which are of interest; these measure the additional benefits of diplomas beyond those of a typical year of education. For movers, D_{12} is .172, indicating a high school diploma raises log wages by .172, i.e., about a 17% increase in wages. This effect is substantively large and statistically significant (standard deviation is only .024). In contrast, a college diploma raises movers' wage by only about 6%, and this effect is not statistically significant. However, in the mature sample it is the college degree which has the greatest effect (about 19%) while high school is worth 8%. Both effects are significant. This pattern reflects the delayed payoff to higher education.

Can the queuing model fit the data? As the "Basic Simulation" column shows, it can. For this simulation, equation (3), which has credential effects, was used for the score equation, and the model was then simulated using the observed sample of people and jobs. The results here will be the baseline for two comparisons, one with a simulation without credential ranking effects, and the second with a simulation with credentials but with equal group sizes.[16]

The "No Credentials" column reports the results of simulation in which equation (4), which gives each year of education the same ranking effect, is used as the score equation. Any credential outcome effects which arise must then be the result of the interaction of group size and the job distribution. Although the simple example above showed that such effects could, in principle, be substantial, they are not here. All credential dummies are small, 1% or less, and the only statistically significant one is negative.

Why aren't the size effects more important? When groups are cleanly separated, as in Figure 3, the size of a group "belongs" to that group alone, in the sense that increasing the size of a group increases the spread of that group but no others. However, when groups overlap, increasing the size of one group also increases the spread of groups which overlap with it. In effect, the size of a group "belongs" to all groups which overlap that group, so group size has no effects which are focussed on the one group. Figure 4 provides an illustration. In the top panel, 11th and 12th graders are separated completely, while in the

[16]Why not compare these later simulations with the actual data? Such comparisons mix two effects: the effects of changes in the assumptions (e.g., no credential ranking effects), and discrepancies caused by the simple functional forms used in the score equations. Comparison with the basic simulation isolates the former, which is of theoretical interest.

bottom panel they are not. The left hand side shows the situation with equal group sizes; the right hand side shows the effects of doubling the size of the 12th grade group. The overall income distribution is the same in all cases, and small tables summarize the effect of each grade on income. Note that when there is no overlap (top panel), increasing a group's size increases the earnings gap between that group and other groups, but has no effect on the gaps between the other groups (i.e., the income gap from grade 11 to 12 and 12 to 13 increases when the size of grade 12 increases, but the gap between 10 and 11 remains unchanged). In contrast, with overlap (bottom panel), increasing the size of the 12th grade has a smaller effect than before, and it also affects the 11th grade.

In short, overlap diminishes the importance of size, since increases in the size of one group affect overlapping groups as well. In this case the effect of each grade is only 1/5 or 1/6 of the standard error (in the ranking procedure), so size doesn't make much difference.

This argument shows that overlap diminishes size effects; however, size may still be important in interaction with credential ranking effects, which increase group separation. To test this, a final set of simulations used equation (3), with credentials, as the score equation, but equalized the size of the different groups. These simulations matched the observed income distribution to a hypothetical sample of people in which there were the same number of people in each grade. This does diminish the credential outcome effects, by a maximum of 13% for mature high school graduates (from .070 to .061). The other effects are more modest, ranging from about 2 to 8%.

In short, these analyses suggest that, for these data, size effects make only minor contributions to observed credentials effects, and that the distinctive value of credentials in the ranking process is critical.[17] While such a conclu-

[17]The safest conclusion is that size effects are small. Either credential ranking effects or the job distribution or their interaction could account for the credential outcome effects. The first of these seems most likely for several reasons. For the income distribution to cause the observed effects, it would have to have an irregular shape which would have a special effect on one group. This is unlikely because overlap is so extensive that this is not possible, and because the equal size simulation caused different grade levels to match up with portions of the income distribution other than those they matched in the original sample. Since the credentials effects persisted, ranking effects seem to be the cause. Additional simulations with artificial job distributions would make this conclusion definitive. This paper concerns size effects, so analysis stops here.

sion supports the usual interpretation of the credentials effects, note that the mechanism is somewhat different. In the usual interpretation, employers value credentials and will pay more for them. If so, a regression model is actually the correct model of the labor market. In this model, in contrast, employers have already fixed the wage for their jobs, but pay particular attention to whether a person has a credential in deciding how to fill the job. This contrast suggests a more extended consideration of regression models, and to that I now turn.

4 A Contrast with Regressions as Models

The dominant theory testing strategy in the social sciences proceeds from some verbal formulation to a specification such as

$$\text{outcome}_i = a_1 X_{1i} + a_2 X_{2i} + \ldots + a_n X_{ni} + \epsilon_i \tag{5}$$

where the X are independent variables, a are coefficients (to be estimated from the data) and ϵ is an error term. The subscript i refers to the i'th unit (person, state, organization). The coefficients receive names in view of the theoretical discussion, and are interpreted in that light. In contrast, the strategy taken above was to translate a verbal model into a mathematical description of process, and then to analyze that process. In this particular case, the choice of strategy makes a difference. This suggests that the usual interpretive weight put on regression results may be excessive.

In a regression-style analysis, the theoretical arguments of this paper become arguments about coefficients on credential dummy variables, as in equation (3). The argument that the size of a group is important would be tested by adding some size variables to the right-hand side of the equation. An appendix describes the results of such tests in detail, but the overall results are easy to summarize:

1. The importance of size is variable and sensitive to the exact specification.
2. Some of the estimated size effects are substantial, much larger than those found in the queuing analysis above.

An irony: the queuing approach suggested that size effects might be important, yet it is the regression analyses, rather than the queuing analyses, which give them the greatest weight.

This paper will not attempt to determine which of these models is more appropriate.[18] However, it seems safe to conclude from this analysis that *if* size effects are important it is not because of the kind of queuing process described above. The goal of this paper is to show that regressions are not the only descriptions of the world, and the choice of model may affect the substantive conclusions one arrives at.

Regressions, if interpreted rigorously as models, make strong assumptions about how the world works. They assume that macro analysis is an aggregation of micro analyses, so that, for example, the structure of the labor market is simply the aggregation of individual outcomes given by regression equations. One can add macroscopic variables to such equations, by making the outcome for an individual depend on the size of different groups, the overall state of the labor market, or the sectoral location of the individual. Yet this is less of a change than it would seem, for even with these additions the whole remains the sum of the parts. Such models predict that changing the mix of people will directly change the mix of outcomes. Much sociological theorizing, for the labor market and for other domains, argues against such a view. We should shape our models accordingly.

5 Conclusion

Many sociological platitudes are remarkably difficult to incorporate into our day to day analysis of the world. This has certainly been the case for the labor market. The theoretical claim that outcomes result from the interplay of the structure of jobs, the demography of people, and the matching process between the two is difficult to translate into empirical analysis; this paper illustrates an admittedly crude way it might be done. The queuing model makes the theoretical distinction between the value of a characteristic in a matching process and its value for outcomes such as income, and shows how demography and the job structure intervene.

The finding that size effects are small is interesting from several points of view. First, it contrasts with the findings of some conventional regression mod-

[18]However, the instability of the regression results indicates their sensitivity to the exact pattern of size and income, making it difficult to put much faith in any particular estimate. In contrast, the queuing model provides not only a consistent finding of minimal size effects but some reasons, related to group overlap, that this should be so.

els, showing that the choice of strategy matters. Second, the size effects were generally positive: bigger groups gained. This contrast with the neoclassical theory again points to the importance of exactly specifying mechanisms.[19]

The queuing model suggested investigating size effects, but the finding that these are small does not invalidate the model. It could not, since the analysis which reached that conclusion assumed the queuing framework, and used it to assign relative weights to size, credentials, and the job structure.

This assumption of the general model may seem troubling; certainly it implies that the results here do not prove that size effects are small, since they do not test the validity of the model. But is this any different from more conventional analyses? Regressions, with controls, find that education affects income and status. At the individual level, investigators assume that, if their controls are good enough, they are measuring how a person's income would change if they had a different level of education. But of course, one can not carry out this experiment, and the conclusion rests on the validity of the model. At the structural level, investigators have asked what the effects on the distribution of income and status would be if the distribution of education were equal, or at least more equal (Chiswick and Mincer, 1972; Jencks and etal, 1972). Their conclusion—that there would be an equalization, but only a small one—depends on the assumption that the market simply aggregates individual outcomes and that the parameters of the regression relationship would remain unchanged as other conditions varied.

Thus routine data analysis rests as much on untested assumptions as the analysis of this paper. Further, the queuing approach here fits the data as well as regression models, so one can not be preferred to the other on grounds of fit. The queuing approach, seems, if anything, more consistent with our theories of the labor market and social structure in general, and with the facts those theories were built on. And, finally, the two approaches yield different conclusions about the same problems. While the queuing model above is an obvious simplification, these considerations suggest that the structural models in general merit further investigation.

[19]The queuing model does *not* predict size effects are positive, it merely indicates that they may be—just as they may be negative given the right conditions.

A Appendix: Regression Estimates of Size Effects

A straight translation from Figure 3 is that the mean income of a group is a function of the total size of all lower educational groups and half the size of that group; call this variable M (for midpoint). This variable can be substituted for the schooling variable, or it can be used in conjunction with years of schooling:

$$\log(w) = \alpha_0 + \beta_1 M + \alpha_3 X + \alpha_4 X^2 + \epsilon \quad (6)$$

$$\begin{aligned} \log(w) &= \alpha_0 + \alpha_1 S + \alpha_2 S^2 + \alpha_3 X + \alpha_4 X^2 + \alpha_5 SX + \\ &\quad \beta_1 M + \epsilon \end{aligned} \quad (7)$$

Alternately, a more conventional test for size effects looks at the size of a group without totalling up the size of all "lower" groups. Letting G denote group size, this leads to

$$\begin{aligned} \log(w) &= \alpha_0 + \alpha_1 S + \alpha_2 S^2 + \alpha_3 X + \alpha_4 X^2 + \alpha_5 SX + \\ &\quad \beta_1 G. \end{aligned} \quad (8)$$

It will be convenient to refer to equation (6) as the "cumulative size only equation," equation (7) as the "cumulative size equation," and equation (8) as the "group size" equation.

If one believed there might be true size and credentials effects, the natural regressions would combine both:

$$\begin{aligned} \log(w) &= \alpha_0 + \alpha_1 S + \alpha_2 S^2 + \alpha_3 X + \alpha_4 X^2 + \alpha_5 SX + \\ &\quad \beta_1 M + \gamma_1 D_8 + \gamma_2 D_{12} + \gamma_3 D_{16} + \epsilon \quad (9) \\ \log(w) &= \alpha_0 + \alpha_1 S + \alpha_2 S^2 + \alpha_3 X + \alpha_4 X^2 + \alpha_5 SX + \\ &\quad \beta_1 G + \gamma_1 D_8 + \gamma_2 D_{12} + \gamma_3 D_{16} + \epsilon. \quad (10) \end{aligned}$$

The β and γ coefficient estimates can be compared with those of previous equations to see how credentials effects and size effects interact.

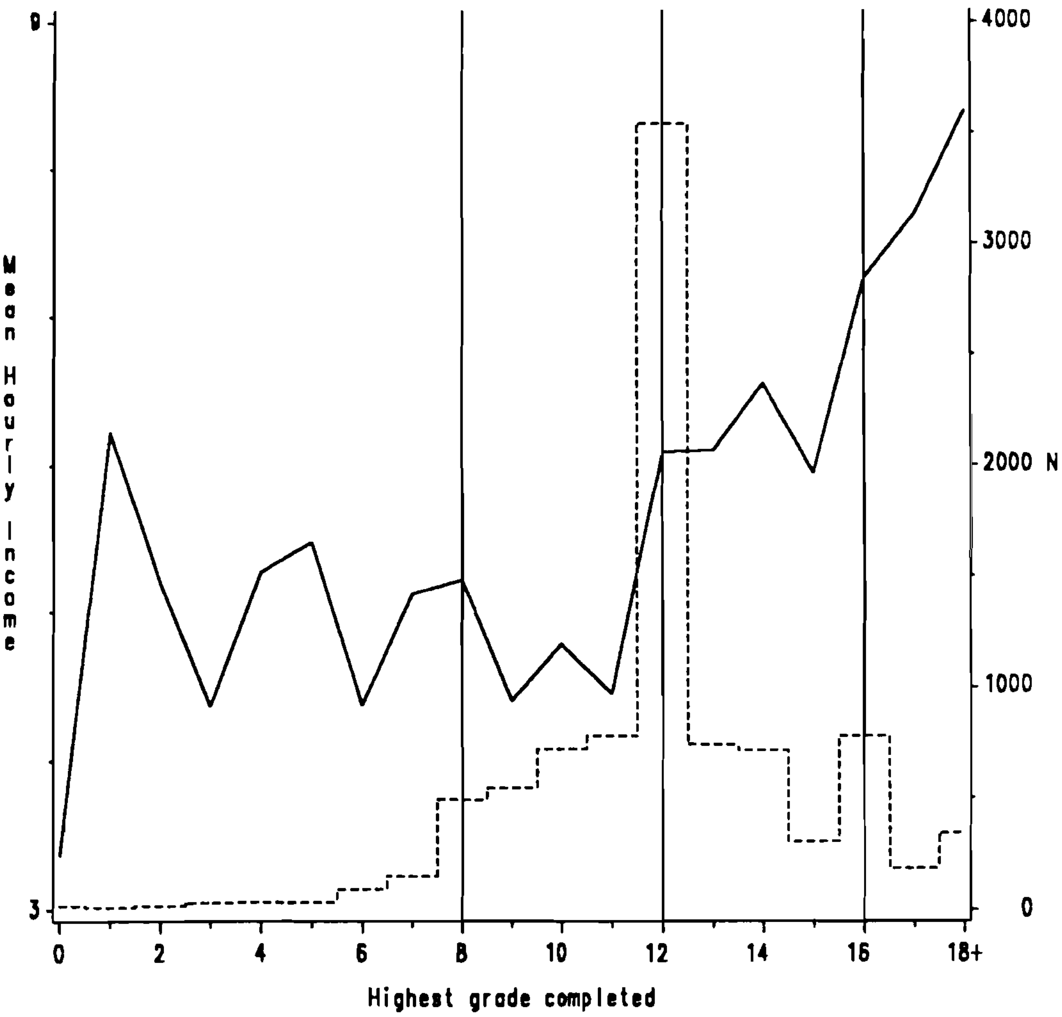
Tables 6 and 7 display the results of these regressions for the mover and mature samples respectively.[20] The dominant impression is one of instability

[20]The size regressions for movers are based on a slightly smaller sample than the main results, size 9,511 rather than 9,560. I believe the discrepancy arises from the exclusion of some low income cases, but will check it. At any rate, all size regressions are carried out on the same sample, so comparisons across regressions should be reasonable.

and contradictory interpretations. For the movers, the size effects are significant and substantively large[21] without controls for credentials, and small and insignificant with controls. This suggests the apparent size effects are spurious. However, some of the regressions with size effects have smaller credential dummies than those without such effects (compare the results for equations (9) and (10) with that for equation (3) for D_{12} and D_{16}), suggesting the estimated credentials effects were spuriously large when size was not considered. Then again, some of the credential dummies increase, and though some of the changes are marked (10% change in estimated credential effect), none are large relative to the errors of the coefficient estimates.

The pattern for mature males is even less clear; the estimated size effect for M doubles when credential dummies are included, and one of the credential effects (D_{12}) is cut in half while the other rises. For this sample, as for the other, the analyses based on cumulative size differ, even in qualitative features, from those based on group size.

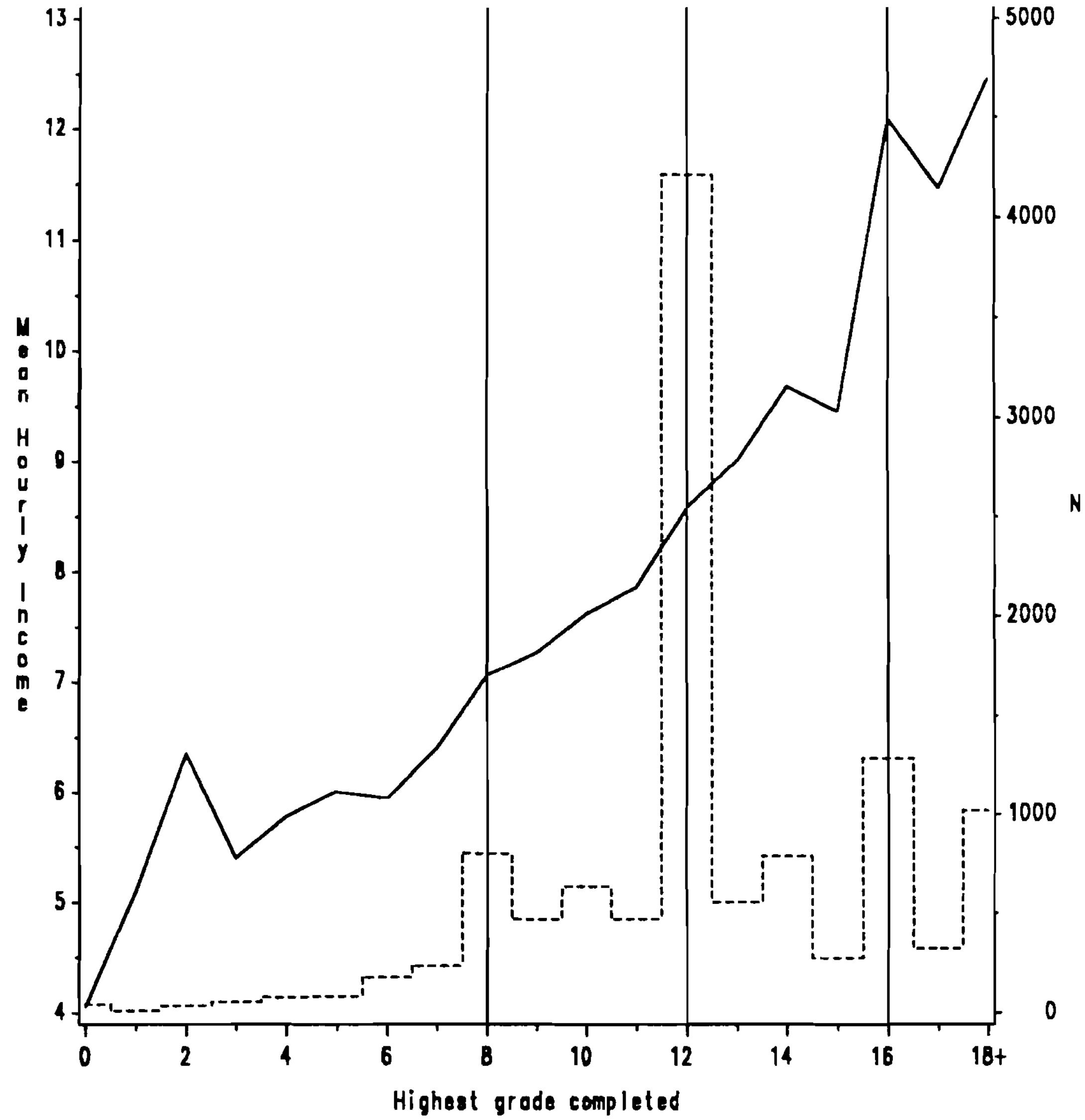
[21] Direct comparison with the simulation results is difficult. However, a rough estimate of the magnitude of size effects comes from multiplying the size coefficients by the observed sizes and then by the sizes which would hold under the hypothetical equal distribution across grade levels. For example, the mover sample has 3534 people in grade 12; an equal distribution of the population would have 771.8 people in that grade, i.e., 2762.2 fewer people. The G coefficient of .034 times this difference (2.76 when units are in 1,000's) is .094. This .094 reduction is large, being 55% of the observed credential effect of .172. Similar calculations can be carried out for other coefficients and samples; changes range from small increases to large (over 100%) reductions in the credentials effects.



The left axis and solid line report income in 1979 dollars per hour.
The right axis and dashed line report number in each grade.
N=9,560 civilians, over age 16, with positive earnings.
Source: March, 1980 Current Population Survey.

Size and Mean Income of Each Grade
White, Male Movers, 1979

Figure 1:



The left axis and solid line report income in 1979 dollars per hour.
The right axis and dashed line report number in each grade.
N=11,571 civilians, aged 40-55, with positive earnings.
Source: March, 1980 Current Population Survey.

Size and Mean Income of Each Grade
Mature, White Males, 1979

Figure 2:

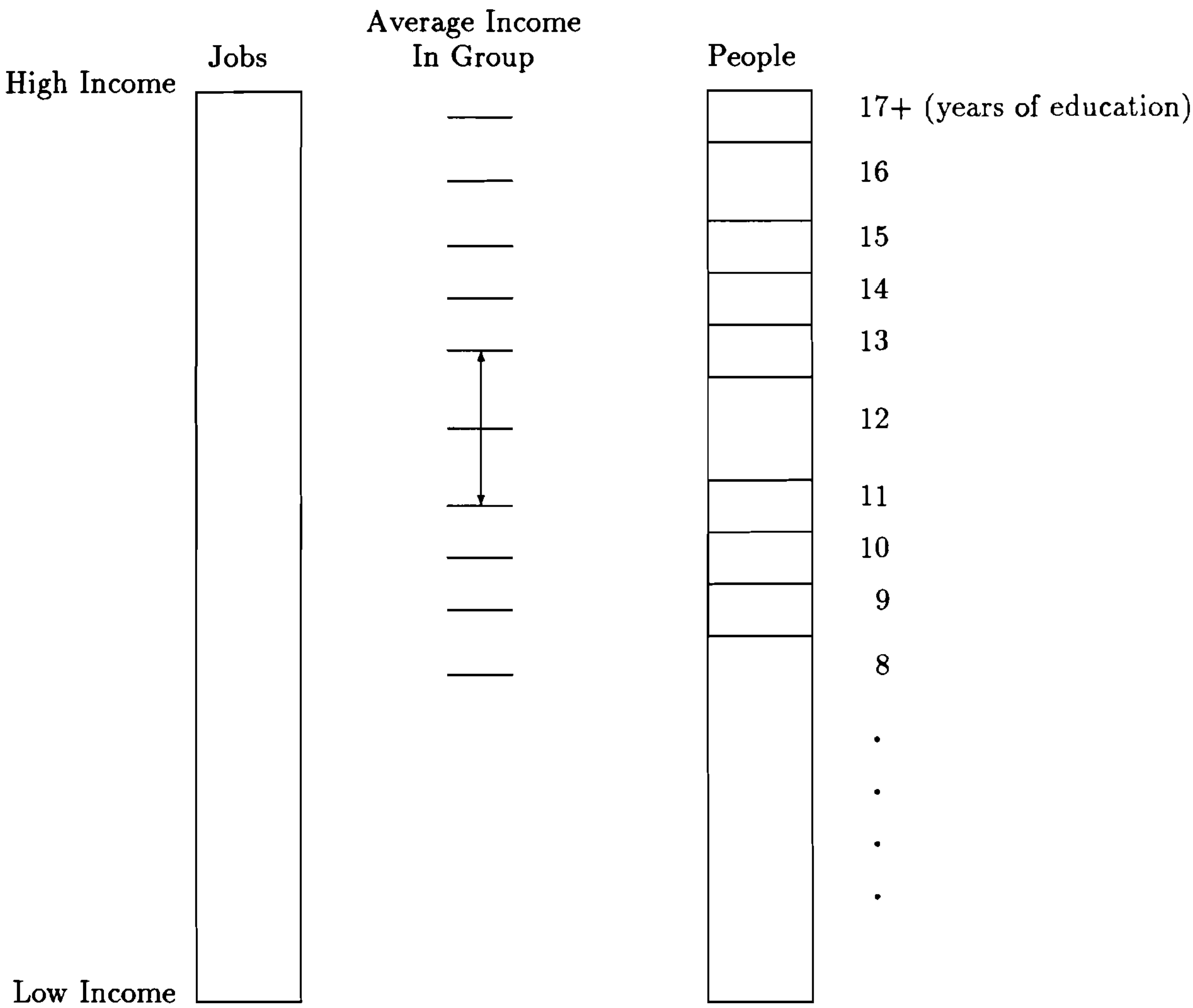


Figure 3: A Queuing Interpretation of Credentials Effects

Person (Grade)	Job (Wage)
13	12.0
12	11.0
12	10.0
11	9.0
11	8.0
10	7.0
10	6.0
9	5.0
9	4.0

Grade	Mean Wage	Delta Wage
13	12.0	1.5
12	10.5	2.0
11	8.5	2.0
10	6.5	

No Overlap: Equal Sizes

Person (Grade)	Job (Wage)
13	12.0
12	11.0
12	10.0
12	9.0
12	8.0
11	7.0
11	6.0
10	5.0
10	4.0

Grade	Mean Wage	Delta Wage
13	12.0	2.5
12	9.5	3.0
11	6.5	2.0
10	4.5	

No Overlap: Unequal Sizes

Person (Grade)	Job (Wage)
13	12.0
12	11.0
11	10.0
12	9.0
10	8.0
11	7.0
10	6.0
9	5.0
9	4.0

Grade	Mean Wage	Delta Wage
13	12.0	2.0
12	10.0	1.5
11	8.5	1.5
10	7.0	

Overlap: Equal Sizes

Person (Grade)	Job (Wage)
13	12.0
12	11.0
12	10.0
11	9.0
12	8.0
12	7.0
10	6.0
11	5.0
10	4.0

Grade	Mean Wage	Delta Wage
13	12.0	3.0
12	9.0	2.0
11	7.0	2.0
10	5.0	

Overlap: Unequal Sizes

Each section presents a matching of persons and jobs with a summary table below it. The table gives the mean wage for each grade and, under "Delta," the difference between the mean wage for each grade and the one preceding it.

Figure 4: Overlap and the Numbers Effect

Table 1: Sample Sizes for March, 1980 CPS

Universe	N
All Records	136,712
Excluding Spanish Supplemental Sample	135,573
Basic Sample*	135,336
White Males	56,558
White Males with Positive Income	45,686
White Male Movers	9,560
White Males aged 40–55	11,571

Each row, except for the last two, is a restriction on the universe of the preceding row. Each of the last two rows is part of the universe of white males with positive income. 1,239 people appear in both of the last two samples.

*The basic sample excludes the Spanish supplemental sample and records with invalid data.

Table 2: Comparison of Samples for Calendar 1979

	SAMPLE*				
	Basic	WM	WM+	WM Mover	WM 40–55
N	135,336	56,558	45,686	9,560	11,571
STATUS (%)					
Didn't work or want work	30%	18%	0%	0%	0%
Self-employed	8	15	17	0	23
One employer and no job search	45	50	62	0	66
Didn't work but searched or was discouraged	1	1	0	0	0
Multiple Employers	11	13	16	75	7
One part-year employer and job search	4	4	5	25	4
Total	100%	100%	100%	100%	100%
DEMOGRAPHICS					
White (%)	88%	100%	100%	100%	100%
Male (%)	47%	100%	100%	100%	100%
Age (years)	41.0	40.4	37.3	29.5	47.3
	(18.8)	(18.3)	(14.9)	(12.1)	(4.6)
Experience (years)		22.3	18.8	11.3	28.8
		(18.8)	(15.2)	(12.4)	(5.7)
Highest Grade Completed	11.6	11.9	12.4	12.1	12.4
	(3.2)	(3.3)	(3.0)	(2.8)	(3.3)
LABOR FORCE DATA					
Hourly Earnings ($/hour)			7.40	5.91	9.09
			(4.6)	(4.0)	(4.8)
Log Hourly Earnings			1.78	1.56	2.03
			(0.76)	(0.70)	(0.68)
Earned Income ($)			14,768	9,334	19,891
			(11,084)	(8,004)	(11,381)
Weeks Worked			45.1	37.8	49.3
			(13.2)	(14.4)	(7.7)
Hours per Week			41.2	39.2	44.4
			(11.5)	(11.3)	(9.0)

Cells report either percentages or means with standard deviations in parentheses. See text for details of variables. Labor force related variables omitted where inappropriate.

*These refer to the samples defined in Table 1 and the text: Basic (all valid data in regular sample), WM (white males), WM+ (white males with positive earnings), WM Mover (white male movers), and WM 40–55 (white males aged 40–55).

Table 3: Percentage of Extreme Values Recoded

	SAMPLE*		
	WM+	WM Mover	WM 40–55
Experience < 0	0.3%	0.7%	0.0%
Hours/week > 70	1.7	1.3	2.4
Hourly Earnings > $24/hr	1.4	0.8	1.9
Hourly Earnings < $.10/hr	0.5	0.1	0.5
N	45,686	9,560	11,571

Variables were recoded to the indicated extreme values if they were more extreme. For example, any cases with experience less than 0 were coded as having experience of 0. The table reports the percentage of cases recoded for each variable in each sample. The recoding intended to limit the influence of single, questionable values on the data analysis. Recoding of hours/week preceded calculation of hourly earnings.

*These refer to the samples defined in Table 1 and the text: WM+ (white males with positive earnings), WM Mover (white male movers), and WM 40–55 (white males aged 40–55).

Table 4: Simulations For Movers

Coefficient	Actual Data	Basic Simulation	No Credentials	Equal Sizes
Intercept	−.036	−.122	.267	−.091
	(.107)	(.009)	(.008)	(.007)
Education	.103	.117	.087	.114
	(.020)	(.002)	(.001)	(.001)
Education^2	−.0016	−.0021	−.0007	−.0020
	(.0008)	(.0001)	(.0001)	(.0001)
Experience	.077	.079	.056	.078
	(.003)	(.000)	(.000)	(.000)
Experience^2	−.0010	−.0010	−.0009	−.0010
	(.0000)	(.0000)	(.0000)	(.0000)
Exp*Eductn	−.0018	−.0020	−.0003	−.0020
	(.0002)	(.0000)	(.0000)	(.0000)
D_8	.106	.102	−.010	.102
	(.051)	(.004)	(.004)	(.003)
D_{12}	.172	.181	.003	.174
	(.024)	(.002)	(.002)	(.002)
D_{16}	.057	.059	.003	.054
	(.040)	(.009)	(.003)	(.002)
Standard Error	.622	.623	.631	.614
		(.000)	(.000)	(.000)

N= 9,560. This is for white, male movers in the 1980 CPS. The last three columns summarize data produced under the indicated assumption. Each row gives the mean and, in parentheses, standard error of the corresponding regression coefficient. Except for the first column, these are derived from 200 simulations, as described in the text.

The "no credentials" simulation assigns scores for ranking based on

$$\text{score} = 0.364 + .069S + .054X - .00089X^2 + \epsilon,$$

where S is years of schooling, X is experience, and the Gaussian error term has standard deviation 0.627. The other simulations use scores based on the coefficients and equation under "Actual Data."

Table 5: Simulations For Mature Males

Coefficient	Actual Data	Basic Simulation	No Credentials	Equal Sizes
Intercept	.422 (.502)	.161 (.041)	.599 (.044)	.281 (.040)
Education	.099 (.030)	.138 (.003)	.119 (.003)	.131 (.003)
Education^2	−.0015 (.0007)	−.0028 (.0001)	−.0018 (.0001)	−.0027 (.0001)
Experience	.039 (.024)	.043 (.002)	.019 (.002)	.041 (.000)
Experience^2	−.0005 (.0007)	−.0005 (.0000)	−.0002 (.0000)	−.0004 (.0000)
Exp*Eductn	−.0006 (.0007)	−.0010 (.0001)	−.0003 (.0001)	−.0008 (.0001)
D_8	.025 (.042)	.020 (.004)	−.007 (.002)	.021 (.002)
D_{12}	.078 (.024)	.070 (.002)	−.010 (.002)	.061 (.002)
D_{16}	.187 (.036)	.161 (.002)	.001 (.002)	.158 (.002)
Standard Error	0.65	.650 (.000)	.649 (.000)	.644 (.000)

N = 11,571. This is for white males aged 40–55 in the 1980 CPS, as described in the text. See Table 4 for further details. The "No Credentials" simulation is based on

$$\text{score} = 0.906 + .070S + .014X - .00015X^2 + \epsilon,$$

where S is years of schooling, X is experience, and the Gaussian error term has standard deviation 0.647.

Table 6: Analysis of the Education-Earnings Relationship for Movers

Coefficient	Equation						
	2	3	6	7	8	9	10
Intercept	−.345 (.094)	−.036 (.107)	.895 (.015)	−.085 (.099)	−.079 (.096)	.029 (.105)	.031 (.104)
Education	.160 (.013)	.103 (.020)		.134 (.014)	.110 (.015)	.100 (.020)	.097 (.020)
Education^2	−.0027 (.0005)	−.0016 (.0008)		−.0034 (.0005)	−.0007 (.0006)	−.0017 (.0010)	−.0014 (.0009)
Experience	.081 (.003)	.077 (.003)	.051 (.001)	.076 (.003)	.077 (.003)	.075 (.003)	.075 (.003)
Experience^2	−.0010 (.0000)	−.0010 (.0000)	−.0008 (.0000)	−.0010 (.0000)	−.0010 (.0000)	−.0010 (.0000)	−.0010 (.0000)
Exp*Eductn	−.0020 (.0002)	−.0018 (.0002)		−.0018 (.0002)	−.0019 (.0002)	−.0018 (.0002)	−.0018 (.0002)
D_8		.106 (.051)				.103 (.050)	.096 (.050)
D_{12}		.172 (.024)				.176 (.030)	.159 (.039)
D_{16}		.057 (.040)				.063 (.042)	.053 (.042)
M (1,000's)			.068 (.002)	.040 (.008)		.001 (.013)	
G (1,000's)					.034 (.005)		.005 (.009)
R^2	.21	.21	.21	.22	.22	.22	.22
Standard Error of Estimate	0.62	0.62	0.59	0.59	0.59	0.59	0.59

Regressions of log hourly earnings on indicated dependent variables. Standard errors reported in parentheses.

G is the size of the educational group; M is the cumulative number of people below a given grade plus half the size of that grade. Both are in units of 1,000.

N= 9,560. This is for white male movers in the 1980 CPS, as described in the text.

Table 7: Analysis of the Education-Earnings Relationship for Mature White Males

Coefficient	Equation						
	2	3	6	7	8	9	10
Intercept	.455 (.503)	.422 (.502)	1.334 (.141)	.415 (.504)	446 (.503)	.387 (.503)	.405 (.503)
Education	.082 (.029)	.099 (.030)		.084 (.024)	.080 (.029)	.104 (.030)	.102 (.030)
Education^2	.0003 (.0005)	−.0015 (.0007)		−.0002 (.0006)	.0004 (.0005)	−.0027 (.0009)	−.0020 (.0008)
Experience	.039 (.024)	.039 (.024)	.016 (.010)	.041 (.024)	.040 (.024)	.041 (.024)	.040 (.024)
Experience^2	−.0005 (.0003)	−.0005 (.0007)	−.0002 (.0002)	−.0005 (.0003)	−.0005 (.0003)	−.0005 (.0003)	−.0005 (.0003)
Exp*Eductn	−.0006 (.0007)	−.0006 (.0007)		−.0007 (.0007)	−.0007 (.0007)	−.0007 (.0007)	−.0007 (.0007)
D_8		.025 (.042)				.039 (.042)	.049 (.044)
D_{12}		.078 (.024)				.036 (.031)	.135 (.040)
D_{16}		.187 (.036)				.205 (.037)	.210 (.038)
M (1,000's)			.070 (.002)	.013 (.009)		.028 (.013)	
G (1,000's)					.004 (.004)		−.012 (.007)
R^2	.10	.10	.10	.10	.10	.10	.10
Standard Error of Estimate	0.65	0.65	0.65	0.65	0.65	0.65	0.65

Regressions of log hourly earnings on indicated dependent variables. Standard errors reported in parentheses.

N = 11,571. This is for white males aged 40–55 in the 1980 CPS, as described in the text.

References

Albelda, R., Gunn, C., and Waller, W., editors (1987). *Alternatives to Economic Orthodoxy: A Reader in Political Economy.* M. E. Sharpe, Inc., Armonk, New York.

Althauser, R. P. (1987). Internal labor markets u.s.a.: a thematic review. Presented at the annual meetings of the ASA, Chicago.

Baron, J. N. (1984). Organizational perspectives on stratification. *Annual Review of Sociology*, 10:37–69.

Beck, E. M., Horan, P. M., and II, C. M. T. (1978). Stratification in a dual economy: a sectoral model of earnings determination. *American Sociological Review*, 43:704–20.

Benavot, A. (1985). *Education and Economic Development in the Modern World.* PhD thesis, Stanford University. Sociology Department.

Berg, I. (1970). *Education and Jobs: The Great Training Robery.* Praeger Publishers, New York.

Berg, I., Finegan, T. A., Bibb, R., and Swafford, M. (1981). Specification in the structural unemployment thesis: issues and prospects. In Berg, I., editor, *Sociological Perspectives on Labor Markets*, chapter 13, pages 347–367, Academic Press, New York.

Blau, P. M. and Duncan, O. D. (1967). *The American Occupational Structure.* Wiley, new york edition.

Blaug, M. (1976). Human capital theory: a slightly jaundiced survey. *Journal of Economic Literature*, 14(3):827–855.

Boudon, R. (1974). *Education, Opportunity, and Social Inequality: changing Prospects in Western Society. Wiley series in Urban Research*, John-Wiley & Sons, New York. A revised version of the 1973 French original.

Boylan, R. D. (1986). A nonparametric, structuralist model of the matching of persons to entry-level jobs. Presented at the Annual Meetings of American Sociological Association in New York City.

Bulow, J. I. and Summers, L. H. (1986). A theory of dual labor markets with application to industrial policy, discrimination, and keynesian unemployment. *Journal of Labor Economics*, 4(3, Part 1):376–414.

Chiswick, B. R. and Mincer, J. (1972). Time-series changes in personal income inequality in the united states from 1939, with projections to 1985. *Journal of Political Economy*, 80(3, part II):S34–S71.

Collins, R. (1971). Functional and conflict theories of educational stratification. *American Sociological Review*, 36:1002–10019.

Collins, R. (1979). *The Credential Society: An Historical Sociology of Education and Stratification.* Academic Press, New York.

Dickens, W. T. and Lang, K. (1984). *A Test of Dual Labor Market Theory.* Working Paper 1314, Cambridge, Ma.: National Bureau of Economic Research.

Featherman, D. L. and Hauser, R. M. (1978). *Opportunity and Change.* Academic Press, New York.

Frank, R. H. (1984). Are workers paid their marginal products? *American Economic Review*, 74:549–71.

Granovetter, M. (1981). Toward a sociological theory of income differences. In Berg, I., editor, *Sociological Perspectives on Labor Markets*, chapter 2, pages 11–48, Academic Press, New York.

Granovetter, M. (1985). Economic action and social structure: the problem of embeddedness. *American Journal of Sociology*, 91(3):481–510.

Heyneman, S. P. and Loxley, W. A. (1983). The effect of primary-school quality on academic achievement across twenty-nine high- and low-income countries. *American Journal of Sociology*, 88(6):1162–1194.

Jencks, C. and etal (1972). *Inequality: a reassessment of the effect of family and schooling in America.* Harper and Row, New York.

Jencks, C. and *et al.* (1979). *Who Gets Ahead?* Basic Books, New York.

Kalleberg, A. L. and Sørensen, A. B. (1979). The sociology of labor markets. *Annual Review of Sociology*, 5:351–379.

Kelley, J. (1973). Causal chain models for the socioeconomic career. *American Sociological Review*, 38:481–493.

Layard, R. and Psacharopoulos, G. (1974). The screening hypothesis and the returns to education. *Journal of Political Economy*, 82(5):985–998.

Lester, R. A. (1954). *Hiring Practices and Labor Competition.* Research Report Series 88, Princeton University, Industrial Relations Section, Department of Economics and Sociology. Also appeared as a book of the same name.

Lockheed, M., Jamison, D., and Lau, L. (1980). Farmer education and farmer efficiency: a survey. In King, T., editor, *Education and Income*, pages 111–152, World Bank, Washington, D.C.

Mare, R. D. and Winship, C. (1984). The paradox of lessening racial inequality and joblessness among black youth: enrollment, enlistment, and employment 1964–1981. *American Sociological Review*, 49:39–55.

Medoff, J. L. and Abraham, K. G. (1981). Are those paid more really more productive? *Journal of Human Resources*, 16:186–216.

Mincer, J. (1974). *Schooling, Experience, and Earnings.* National Bureau of Economic Research, New York.

Piore, M. J., editor (1979). *Unemployment and Inflation: Institutionalist and Structuralist Views.* M. E. Sharpe, Inc, White Plains, N.Y.

Rosen, S. (1977). Human capital: a survey of empirical research. In Ehrenberg, editor, *Research in Labor Economics.*

Sørensen, A. B. (1977). The structure of inequality and the process of attainment. *American Sociological Review*, 42(6):965–978.

Spence, M. A. (1974). *Market Signaling: The Information Structure of Hiring and Related Processes.* Volume 143 of *Economic Studies* ., Harvard University Press, Cambridge.

Stiglitz (1984). *Theories of Wage Rigidity.* Working Paper Series 1442, NBER, Cambridge, Ma.

Stolzenberg, R. M. (1978). Bringing the boss back in: employer size, employee schooling, and socioeconomic achievement. *American Sociological Review*, 43:813–828.

Thurow, L. C. (1968). Disequilibrium and the marginal productivity of capital and labor. *restat*, 50(1):23–31.

Thurow, L. C. (1975). *Generating Inequality: Mechanisms of Distribution in the U.S. Economy.* Basic Books, New York.

Thurow, L. C. (1983). *Dangerous Currents: The State of Economics.* Random House, New York.

U.S. Dept. of Commerce, Bureau of the Census (1984). *Current Population Survey: Annual Demographic File, 1980 (with 1980-Based Weights).* Inter-university Consortium for Political and Social Research, P.O. Box 1248, Ann Arbor, Michigan 48106, second icpsr edition. ICPSR 8040.

White, H. (1970). *Chains of Opportunity: System Models of Mobility in Organizations.* Harvard University Press, Cambridge, Ma.

Yellen, J. (1984). Efficiency wage models of unemployment. *American Economic Review*, 74(2):200–205.

Recent Working Papers Produced in IIASA's Population Program

Copies may be obtained at a cost of US $ 5.00 each from IIASA's Publications Department.

WP-87-51, *The Concentration of Reproduction: A Global Perspective* by W. Lutz. June 1987.

WP-87-58, *A Simple Model for the Statistical Analysis of Large Arrays of Mortality Data: Rectangular vs. Diagonal Structure* by J. Wilmoth and G. Caselli. June 1987.

WP-87-59, *Sibling Dependences in Branching Populations* by P. Broberg. June 1987.

WP-87-87, *The Living Arrangements and Familial Contacts of the Elderly in Japan* by K. Hirosima. September 1987.

WP-87-92, *The Demographic Discontinuity of the 1940s* by N. Keyfitz. September 1987.

WP-87-104, *A Random-Effects Logit Model for Panel Data* by D. Wolf. October 1987.

WP-87-116, *Some Demographic Aspects of Aging in the German Democratic Republic* by T. Büttner, W. Lutz, and W. Speigner. November 1987.

WP-88-10, *On the Concentration of Childbearing in China, 1955–1981* by W. Lutz. February 1988.

WP-88-13, *Beyond "The Average American Family": U.S. Cohort Parity Distributions and Fertility Concentration* by M. King and W. Lutz. March 1988.

WP-88-23, *Understanding Medical and Demographic Trends with MEDDAS* by M. Rusnak and S. Scherbov. April 1988.

WP-88-32, *Kinship Patterns and Household Composition of the Elderly: Hungarian Women, 1984* by D. Wolf. April 1988.

WP-88-36, *"DIAL" - A System for Modeling Multidimensional Demographic Processes* by S. Scherbov and V. Grechucha. May 1988.

WP-88-44, *Kin Availability and the Living Arrangements of Older Unmarried Women: Canada, 1985* by D. Wolf, T. Burch, and B. Matthews. June 1988.

WP-88-46, *Population Futures for Europe: An Analysis of Alternative Scenarios*, by D. Wolf, B. Wils, W. Lutz, and S. Scherbov. June 1988.

WP-88-90, *Comparative analysis of Completed Parity Distributions: A Global WFS-Perspective*, by W. Lutz. October 1988.

WP-88-104, *Future Regional Population Patterns in the Soviet Union: Scenarios to the Year 2050*, by S. Scherbov and W. Lutz. November 1988.

WP-88-120, *AIDS and HIV Surveillance in Europe*, by M. Artzrouni and G. Heilig. December 1988.

WP-88-124, *DDMSLT: A Computer Program for Estimating the Duration-Dependent Multistate Life Table Model*, by C. Calhoun. December 1988.

WP-89-05, *Multi-State Analysis of Family Dynamics in Austria: Scenarios to the Year 2030*, by W. Lutz and S. Scherbov. January 1989.

WP-89-06, *The Demographic Dimensions of Divorce: The Case of Finland*, by W. Lutz, B. Wils, and M. Nieminen.