ALTERNATIVE MODELS OF EARNINGS DETERMINATION AND LABOR MARKET STRUCTURES*

ERIC A. HANUSHEK

ABSTRACT

There are three distinct research traditions in the analysis of individual earnings determination: human capital, or earnings function, analyses; aggregate wage analyses; and labor demand analyses. An important and incongruous aspect of each is the treatment of geographical differences in labor markets. This paper first investigates the magnitude and character of geographical wage differentials. The sizable differences discovered there are then related to the existing, and highly simplified, models of labor market differences. While the two major classes of models (compensating differentials and labor demand) differ signficantly in assumptions and implications, it is impossible to distinguish adequately between them. There appears to be a clear need for more structural analyses of labor market operations.

Research into the structure of individual earnings has been voluminous. Most recent analyses, under the heading "human capital" analysis, concentrate on "quality" differences among workers. Nevertheless, other research traditions addressing essentially the same questions have taken quite different, and conflicting, views of wage determination. These include investigations of aggregate differences in earnings patterns arising from differences in employing industries, in occupations, and in employment location and analyses based upon production relationships and the derived demand for labor.

A major difference among the alternatives relates to labor market definition and the modeling of how labor market structure affects indi-

The author is Professor of Economics and Political Science and Director, Public Policy Analysis, University of Rochester.

* Helpful comments on previous versions were provided by Robert Evenson, Gary Fields, Richard Nelson, Randall Olsen, Walter Oi, John Quigley, Paul Schultz, and Kenneth Wolpin [Manuscript received April 1979; accepted September 1979.]

vidual earnings. Human capital research concentrates upon differences among individual workers while generally assuming all workers participate in a common aggregate labor market. On the other hand, direct analyses of aggregate labor market differences (denominated by the geographic area, industry, or occupation of workers) display sizable differences across labor markets but generally ignore differences among individual workers. Finally, demand studies, concentrating on differences in labor and market structure, typically ignore differences and responses of individual workers. While the different classes of research have proceeded quite independently, available evidence suggests that each has a role in explaining individual earnings.

The first section briefly reviews these major research traditions with an emphasis upon the conflicting assumptions and evidence of each. The second section, which melds the different approaches, presents new empirical evidence about the importance of geographic differences in earnings. These findings have implications both for research into earnings determination and for assessment of a wide range of programs (such as income support or economic development) that depend importantly upon the interpretation and understanding of the processes generating such differences. The third section assesses the existing, although limited, set of models specifically considering geographical earnings variations. This assessment is facilitated by the more precise estimates of underlying differences from Section II, and for the first time a direct comparison of the alternative regional models is possible. In the end, conceptual shortcomings rather than data limitations appear to be most important. The underlying themes are: (1) that increased attention should be devoted to the underlying structural relationships, and (2) that newly available micro-data for geographically separated, albeit interdependent, local labor markets considerably enhance our ability to investigate such structural relationships.

I. PAST RESEARCH

Past research into regional wage differences, while chiefly descriptive, indicates large and persistent wage differences across whatever level of regional aggregation is used.[1] Nevertheless, quality differences among individuals are seldom considered,[2] and few insights are provided about causes of observed differences.

1 See, for example, Bloch [2], Hanna [18, 19], NBER [34], Segal [42], Gallaway [13], Fuchs [11], and Chiswick [5]. The level of aggregation is often dictated by the data and not chosen by the researcher. Past analyses have concentrated upon the differences among Census regions or states.

2 Exceptions include Fuchs [11], Chiswick [5], and Hirsch [26].

Human capital research, on the other hand, places emphasis on investments by individuals (e.g., schooling and on-the-job training) and the differential expected earnings related to such human capital investments (see reviews by Mincer [32], Blaug [1], and Rosen [38]). From this research, the relationship between earnings and schooling and training investments is quite commonly accepted. However, consideration of specific empirical results gives some pause. Individual studies—based upon a variety of data sources, many of which are rather specialized—have emphasized different aspects of earnings determination and, thus, direct comparisons of empirical results are difficult. Nevertheless, allowing for such differences, estimates of rates of return to different investments show wide variation.[3] Estimated rates of return for years of schooling, particularly in regression estimates considering other individual differences, appear very unstable: Changes in sample, changes in time periods, and changes in precise model specifications yield enormous changes in estimated rates of return. Attempts to improve these models by adding more detailed measures of individuals (say, ability or school quality) or characteristics of labor markets (see Section III, below) have not narrowed the coefficient differences significantly (see also Blaug [1]).

Interpretive difficulties with these investigations arise from two sources. First, the conceptual models involve purely supply-side behavior of individuals, but the empirical models are actually complicated reduced-form relationships that combine supply and demand forces. Second, there is little consensus on the appropriate specification of the underlying structural relationships.

The final strand of research, starting from a very different perspective, takes the supply of workers as exogenous and concentrates on demand relationships. These studies start from aggregate production functions and investigate the derived demand for human capital from a perspective of labor substitutability. The related empirical studies, based upon aggregate time-series or international comparisons, have yielded imprecise, and often inconsistent, results.[4]

---

3 "Rate of return" is used as a shorthand description of increases in earnings associated with different investments. With a series of assumptions (cf. Mincer [33]), the proportionate increase in earnings associated with an additional year of schooling can be interpreted as the rate of return to added schooling. Comparisons of rates of return to schooling or other investments are in many ways similar to comparisons of average wage differentials in previous aggregate studies. A "rate of return" interpretation of estimated school-earnings relationships for local areas (Section II below) is, of course, much more difficult since it requires additional assumptions about lifetime mobility.

4 Bowles [4], Dougherty [6], Fallon and Layard [9], Griliches [16], Dresch [7], and Johnson [28]. This research has been extended by Nelson and Phelps [35] and Welch [49] to include dynamic factors and technological change.

One seldom-discussed issue—the appropriate definition of labor markets—is common to the human capital and the demand analysis and is central to the analysis here. This paper concentrates upon geographical aspects of labor market definition. While diverging significantly in basic approach, human capital and demand analyses share a common assumption that national samples are appropriate for empirical work.[5] This assumption, while often an empirical necessity because of data availability, in part reflects simple theoretical arguments: Free movement of factors of production or, even with barriers to factor movement, free movement of goods (according to factor price equalization theorems) suggests equalization of relative factor payments. Evidence that at least money capital prices are roughly equilibrated across regions (Straszheim [44]) then implies absolute labor prices should be equalized. This suggests that the entire country can be viewed as a single labor market with any observed earnings differentials simply reflecting temporary phenomena or statistical artifacts.

On the other hand, there is reason for skepticism. The theoretical results are static equilibrium statements (indicating nothing about the path or speed of adjustment) and are derived from strong assumptions. To the extent that there are barriers to resource movement (such as transportation costs), important nontraded goods (such as services), differences in production functions across markets, or economic or population growth dynamics which counteract migration adjustments, the predicted static equilibriums may be obtained only after a long period of time, if at all.

Empirical evidence, beginning with aggregate regional wage studies that consistently show large regional differentials, also consistently indicates distinct labor submarkets. Further, virtually every micro-data human capital study of schooling-earnings relationships that allows regional variation (through regional dummy variables or stratification) finds significant differences: for aggregate U.S. regions (e.g., Hanoch [20]); for states (e.g., Chiswick [5], Welch [48], Smith and Welch [43]); and even for separate metropolitan areas (e.g., Hanushek [21, 23], Hirsch [26]).

Submarket differences introduce questions about the interpretation of past earnings functions and demand relationships based upon national or macro region samples. However, an initial consideration is the persistence of observed cross-sectional variations since, if observed differences simply reflect transitory local phenomena, the empirical models still can be inter-

5 The human capital studies are generally based upon individual data for all individuals in a national (or large macro-region) sample, while the demand studies usually use aggregate data for a nation or several nations. (Exceptions to this in the demand studies include Dougherty [6] and Johnson [28] which rely upon cross-sectional U.S. state data.) In addition to questions about the use of national data (above), the demand studies are subject to further questions about the homogeneity of production technologies and labor market structure across either nations or long time periods.

preted as providing reasonable average or long-run estimates.[6] Interestingly, past research has incorporated quite incongruous interpretations of regional differences: at times they are considered insignificant or transitory (as in past earnings function estimation) and at other times significant and persistent enough to elicit long-run investment responses (as through individual migration or production decisions by firms). Direct investigation of the intertemporal persistence of earnings differentials standardized for individual quality differences (and parallel to the cross-sectional evidence below) is not possible because of inadequate historical data. However, a variety of indirect evidence provides a strong prima facie case of the persistence of geographical differentials.

Aggregate analyses of mean wages show at best modest narrowing of regional differences (Borts [3], Segal [42], Fuchs and Perlman [12], and Gallaway [13]). Important additional evidence comes from examining patterns of internal migration. If we accept the prevailing view that earnings differentials are an important determinant of migration patterns,[7] then stability in the levels and patterns of migration strongly suggests stability of earnings differentials. As shown in Table 1, migration flows are remarkably consistent in recent decades. Correlations of state net migration rates for adjoining decades are over .85 and for the decades of the 1940s and 1960s are .77. Also, median state income (not corrected for compositional differences) is correlated .95 in adjoining decades and .87 across two decades, even though there is some decline in the coefficient of variation (from .21 to .16) between 1950 and 1970. On the industry, or demand, side, the correlations of percentage of state employment in manufacturing are .98 for 1950–60, .96 for 1960–70, and .90 for 1950–70. Finally, as shown by Tideman [47] and others, persistent differences exist among metropolitan areas in local employment rates. While these aggregate statistics may mask important compositional differences, the overall picture is one of significant and persistent differences in local labor markets. Even though net migration rates have been large (ranging from −15 percent to +50 percent at the state level between 1960 and 1970), there has been only modest adjustment in area earnings patterns.

Local labor markets clearly exhibit sizable variations in structural aspects—in educational and age distribution of the labor force, in demand

6 One interpretation of past estimates (considered below) is that they reflect average parameters where the underlying data are generated by a random coefficient model. However, persistent labor market effects on the parameters would imply that they are drawn from distributions with different means, and OLS estimation would be inconsistent.

7 Most migration research (see Greenwood [15]) takes this view of migration. An alternative "equilibrium" view is that migration represents life-cycle changes in the evaluation of earnings and amenities of areas.

TABLE 1
SIMPLE CORRELATIONS OF NET MIGRATION RATES AND MEDIAN FAMILY INCOME FOR 48 CONTINGUOUS STATES, 1940–1970

	Net Migration 1940–50	Net Migration 1950–60	Net Migration 1960–70	Median Income 1950	Median Income 1960	Median Income 1970
Net mig. 40–50	1.00					
Net mig. 50–60	.87	1.00				
Net mig. 60–70	.77	.86	1.00			
Med. income 50	.56	.41	.33	1.00		
Med. income 60	.66	.54	.48	.95	1.00	
Med. income 70	.61	.49	.51	.87	.95	1.00

Source: U.S. Bureau of the Census. *Historical Statistics of the United States, Colonial Times to 1970* (Washington: U.S. Government Printing Office, 1975).

for labor and in unemployment rates at any point in time, in industrial and occupational composition of employment, and in the attractiveness or amenities of areas. Therefore, it should not be surprising that these structural factors are reflected in the structure of earnings across labor markets. This is documented in the next section.

II. REDUCED-FORM EARNINGS MODELS IN LOCAL LABOR MARKETS

Consider a reduced-form relationship relating individual earnings to individual characteristics (such as schooling and training). With some generality, this can be represented as a random coefficients model such as:

$$y_{ij} = X_i\beta_{ij} + \epsilon_i \tag{1}$$

where y_{ij} is earnings of individual i in labor market j, X_i is a vector of individual attributes, β_{ij} is a vector of returns to the individual attributes, and ϵ_i is a stochastic term. The expected returns to individual attributes, consistent with a "hedonic" interpretation, are in turn related to aggregate supply and demand characteristics of the labor market such that

$$E(\beta_{ij}) = \bar{\beta}_j = L_j\Lambda + \omega_j \tag{2}$$

where the mean return is constant within a labor market but is a function of labor market characteristics, L_j. This model follows in the spirit of past earnings function estimation and highlights the importance of labor market definition. If the relevant labor market is a national one, L_j is constant for individuals; this implies $\bar{\beta}_j$ is constant for all, and OLS applied to equation

(1) provides consistent estimates of the mean reduced-form returns (or hedonic prices) for individual characteristics. However, if the relevant labor market is local and L_j varies across geographical areas (because of different supplies and demands of individual characteristics), estimation of equation (1) across labor markets no longer provides consistent estimates of the reduced-form parameters of earnings.

Some recent earnings analyses, discussed in more detail below, in fact concentrate on the influence of characteristics of local areas, L_j (e.g., Fuchs [11], Thurow [46], Hall [17], Rosen [39]). However, the approach in these has been to estimate models such as:

$$y_{ij} = X_i\bar{\beta} + L_j\Lambda^* + \nu_i \qquad (3)$$

This "gross standardization," that constrains the returns to individual characteristics to be a constant ($\bar{\beta}$), does not in general provide consistent estimates of either hedonic prices for individual characteristics or the reduced-form parameters for labor market factors, Λ.[8] The severity of problems with the estimation of equation (3) is related directly to the variance of the $\bar{\beta}_j$s.

The remainder of this section concentrates upon the estimation of $\bar{\beta}_j$ for different local labor markets. The following section uses these estimates to investigate equation (2), the relationship between returns to individual factors and labor market characteristics.

Basic Models of Individual Earnings

This analysis integrates the perspective of aggregate earnings analyses, which have concentrated upon geographical variations but have generally neglected compositional differences in the labor force, with earnings function analyses, which describe returns to individual characteristics but generally ignore geographical variations. The empirical strategy involves two steps: (1) dividing the country into distinct local labor markets (defined as SMSAs or County Groups if not in an SMSA), and (2) estimating earnings functions for each market using data from the 1970 Census Public Use Sample.[9] This paper concentrates upon earnings of white males, although it is part of a larger analysis that considers other race/sex groups.

---

8 The key question is whether characteristics of local economies affect the returns to individual factors. Most earnings models are actually semilogarithmic; in these, if the returns are proportional to characteristics of local economies, estimation of equation (3) (where the dependent variable is $\ln y_{ij}$) is acceptable. However, as the analysis below suggests, these returns are not simply proportional to aggregate characterisitics, L_j.

9 The appropriate geographical definition of a labor market has seldom been considered. This analysis, while far less constrained than previous analyses in areal definition, is still confined by available data. The 1/100 Public Use Sample identifies geographical areas comprised by continguous counties with 250,000 or more people and, where possible,

Within each labor market (j), separate models specified as equation (4) are estimated for two schooling groups (high school or less and more than high school).[10]

$$\ln Y_{ij} = \beta_{1j} + \beta_{2j}S_i + \beta_{3j}E_i + \beta_{4j}E_i^2 + \epsilon_i \tag{4}$$

where $\ln Y$ = logarithm of earnings; S = years of schooling, E = years of potential experience (Age − S − 6); ϵ = stochastic element; and β_{1j}, β_{2j}, β_{3j}, β_{4j} = parameters to be estimated.

This paper presents estimation results for prime-age white males who were full-time, full-year workers.[11] Overall, there are 341 regions of the country: 147 SMSAs and 194 separate County Groups, or non-SMSA, areas. A total of 259,894 observations are used to estimate 682 separate earnings models (for the different labor market and school group stratifications).

Estimation Results

This estimation indicates a complicated pattern of individual earnings determination. An overall decomposition of earnings variations, consistent with the stratification and estimation described, is presented in Table 2.[12] For white males with high school or less schooling, 8 percent of total earnings variation arises from differences in mean labor market earnings with less variation across SMSAs than non-SMSAs. Further, consistent with the hypothesis that the market for college-educated workers is more

follows SMSA definitions; country groups within a single SMSA were aggregated. A problem exists when the appropriate labor market definition is smaller than can be identified with the data; if larger areas are more appropriate, this should be apparent in the earnings function estimation, except for sampling errors. The problems appear most serious in the non-SMSA County Groups, which can be large heterogeneous areas, and the subsequent analysis concentrates on SMSAs.

10 For a discussion of the empirical specification, see Mincer [33]. Labor market experience is not observed; instead potential experience (time out of school) is estimated and used. Stratification by schooling levels allows returns to vary with schooling and allows for potential differences in the appropriate definition of labor markets (cf. Hanoch [20]).

11 Full-time, full-year workers were used to separate cyclic phenomena from underlying structural differences. Past evidence suggests that hours worked are more sensitive to cyclic conditions than wages, implying that wage rates are a better indication of longer run conditions than are annual earnings. However, the Census data do not include wage rates or annual hours. For full-time (35+ hours), full-year (48–52 weeks) workers, annual earnings are approximately a linear transformation of wage rates. Nevertheless, if workers with different amounts of labor force attachment receive systematically different wages (cf. Mincer [33]), the earnings estimates may not be representative of the entire labor force. The possibility that more cyclically sensitive regions have higher overall wage rates is considered explicitly in the next section.

12 Similar decompositions by Census regions, along with the sample stratifications of labor markets and individuals, are available from the author.

TABLE 2
DECOMPOSITION OF VARIANCE IN LOG EARNINGS
(White Males, Full-Time/Full-Year Workers)

Group	Between Labor Markets	Explained Within Labor Markets	Total Explained
Schooling ≤ 12 years			
Total	.079	.101	.172
SMSA	.038	.105	.139
Non-SMSA	.051	.094	.140
Schooling > 12 years			
Total	.051	.200	.241
SMSA	.027	.209	.230
Non-SMSA	.026	.177	.198

national and that their higher mobility rates lessen geographic labor market differences, the between-labor-market variation is smaller (5 percent) for college-educated workers. Importantly, considerable variation among individual labor markets exists even within Census regions for both schooling classes.

Schooling and experience differences among workers consistently explain more within-labor-market earnings variation among the college-educated (20 percent for the nation) than among the less educated (10 percent for the nation).[13] The total variation explained by labor market stratification and by individual characteristics, 17 percent for the less educated group and 24 percent for the more educated group, is quite comparable to previous investigations of individual earnings.[14]

Underlying the aggregate decomposition of earnings variations is an interesting picture of the interaction of individual characteristics and the character of labor markets. The means and standard deviations of the

13 Comparisons of explained variation between schooling stratification warrant caution. However, since the lower schooling stratification is likely to be more heterogeneous, one would generally expect variance explained by the measured characteristics to be higher there than in the more schooled samples which it is not.

14 The explained variance from the earnings models, while modest, is not particularly surprising given the stratification not only by labor force status (full-time/full-year workers) but also by schooling, geographical area, race, and sex. Similar earnings decompositions for blacks and females indicate: (1) between-region variance is consistently higher for the less educated; (2) between-region variance is, except for more educated blacks, larger than for comparable whites; and (3) within-region explained variance for females is only slightly less than for males even given the significant errors in the experience estimation for females.

TABLE 3
SCHOOLING COEFFICIENTS—WEIGHTED MEANS[a]
(Standard deviations in parentheses)

Census Region	Schooling ≤ 12 Years SMSA	Schooling ≤ 12 Years Non-SMSA	Schooling > 12 Years SMSA	Schooling > 12 Years Non-SMSA
New England	.044	.044	.101	.096
	(.007)	(.009)	(.011)	(.032)
Middle Atlantic	.051	.056	.111	.105
	(.013)	(.017)	(.018)	(.035)
East North Central	.047	.053	.095	.084
	(.010)	(.020)	(.016)	(.035)
West North Central	.058	.056	.103	.096
	(.014)	(.023)	(.014)	(.029)
South Atlantic	.059	.059	.111	.104
	(.017)	(.017)	(.020)	(.035)
East South Central	.068	.074	.097	.095
	(.017)	(.024)	(.022)	(.040)
West South Central	.047	.056	.101	.091
	(.016)	(.019)	(.021)	(.034)
Mountain	.048	.038	.112	.097
	(.013)	(.023)	(.016)	(.026)
Pacific	.031	.046	.098	.079
	(.013)	(.027)	(.017)	(.023)
All	.049	.056	.103	.093
	(.016)	(.022)	(.019)	(.034)

a Means and standard deviations are weighted by the number of observations in the individual labor market regressions. Since the Public Use Sample is a random sample of the population, this is equivalent to weighting by the population in the relevant group.

school coefficients in Table 3, broken down by schooling group and by Census regions, display part of this interaction.[15] Several aspects of the results are noteworthy. First, mean earnings gains from additional schooling, both in the aggregate and in separate Census regions, are higher for the more educated (10 percent) than for the less educated group (5 percent). Second, schooling coefficients show substantial variation. For example, the mean return (at the Census region level) to a year of schooling ranges from 3 to 7 percent for the less educated group residing in SMSAs;

15 The schooling coefficient is roughly the proportional increase in earnings associated with an additional year of schooling and, with assumptions about the direct costs of schooling, is interpreted as approximately the private rate of return to a year of school; see Mincer [33]. Here, such interpretation also requires an assumption that individuals remain in the same region.

the standard deviation across all SMSAs for this group is 1.6 percent and is even higher for other groups.[16] Third, both differences in Census region means and coefficients of variation (within Census regions and the entire country) consistently indicate less variation in returns to schooling for the higher education group than for the lower education group.[17] Thus, differences in mean earnings across labor markets and differences in the shape of earnings profiles within regions support the hypothesis that earnings variations are less for the more mobile (more educated) individuals.

Similar variations (not shown) are observed for the experience parameters, although the variation appears quantitatively smaller. In general, earnings profiles for the more educated tend to be more peaked than for the less educated (i.e., the linear term is larger but the quadratic term is more negative for the higher schooling group).[18] The important point is again that significant interactions between the shape of the earnings profile and local labor markets are observed.[19]

---

16 Clearly, some variation arises from sampling errors. The standard deviations of coefficients, even within Census regions, are uniformly greater than the estimated standard errors for the separate labor market coefficients. While the standard covariance test is not very powerful here, the hypothesis of homogeneity within Census regions was rejected at standard significance levels (available from author). Since regions have individuals that systematically differ in the location of schooling (because of historic migration patterns), some variation may simply reflect differences in average school quality. Previous attempts to consider quality-equivalent years of schooling produced results which were indistinguishable on statistical grounds from the analysis of quantity of schooling; see Hanushek [23]. An alternative, albeit equally as crude, attempt to consider school quality introduced dummy variables, reflecting Census region of birth, into each local earnings model, but the estimated effects were generally insignificant and showed no consistent pattern.

17 Each of the above comparisons also holds for females and blacks. Schooling coefficients are very similar for males and females and lower for blacks in the less-educated category when compared to whites of the same sex and schooling category. The coefficients of variation are consistently higher for blacks and females than for white males.

18 An alternative interpretation of the experience terms, proposed by Welch [50] is that they reflect quality differences—or vintage effects—of schooling. While vintage effects and "investment" effects are not separately identified, a strong vintage interpretation would imply that the estimated experience parameters would differ across regions because of differences in regional mixes of schooling location, even if the returns to standard quality schooling were the same across regions. The estimated schooling parameter would then be the returns to the average quality of schooling within the given local labor markets, which again could vary by regions even if the return to quality-standardized schooling were in fact the same across regions. Nonetheless, both the direct (but crude) tests of school quality effects (fn. 16) and the magnitude of differences within Census regions (where differences in average school quality might be presumed to be rather modest) for the homogeneous grouping of white males suggest more fundamental labor market differences than simply quality differences.

19 While the estimated schooling coefficients for males and females were very similar, the average experience coefficients (by schooling level and SMSA/non-SMSA) for females

Complete description of variations in earnings relationships across labor markets is complicated, since earning opportunities depend jointly upon the schooling level, experience, sex, race, and geographic location of the individual. To summarize the combined effects of these factors, expected present values of lifetime earnings within labor markets (based upon the estimated earnings functions) were calculated for persons with 12 and 16 years of schooling. Even aggregated to Census regions, the range of estimated lifetime earnings is $21,000 for college graduates and $26,000 for high school graduates (12 and 18 percent of the average, respectively). Moreover, the best labor markets for college graduates are not necessarily the best for high school graduates; that is, labor markets are not "good" or "bad," independent of the characteristics of individuals.

In summary, significant earnings differences—both in the level and shape of earnings profiles—exist across local labor markets. These findings cast considerable doubt on the assumption of homogenous aggregate labor markets common to most past earnings analysis. They also highlight the need for understanding better the operations of local labor markets. A wide variety of public programs—concerned with economic development, income support, unemployment, etc., for local areas—rely upon particular interpretations of local differences and specific assumptions about the operation of local markets. These assumptions often differ by program and are generally unsupported by any analysis. An important by-product of this analysis is the ability to obtain estimates of the underlying geographical earnings differentials, more appropriately standardized for individual characteristics, that can be used in assessing models of geographical earnings differences.

III. SIMPLE MODELS OF REGIONAL DIFFERENCES

Consider the important behavior involved in the interactions among local areas in an open economy. Individuals and firms, who may differ in evaluations of areas, are free to move and are generally presumed to respond to local differences in wages and other attributes. At the same time, exogenous factors such as changes in product demand, demographic shifts, and technological change—affecting industries differently—will, at any point in time, have quite different geographical impacts. Finally,

---

range between .43 and .48 of those for males. This is very close to the average difference in employment rates; in 1970 the ratio of female to male employment rate (employed/total population) was .52. There are, of course, a number of reasons for caution here—employment rates vary by time period and point in the life cycle, and therefore the aggregates do not represent the expected employment for each individual in a given year (particularly when considering full-time/full-year workers).

relocation costs and externalities reflecting historical location patterns will impede the adjustments of both firms and individuals.

Jointly modeling the interactions of these factors is obviously difficult, and the complexity of the task, coupled with a lack of suitable data, has led to the consideration of only highly simplified models. Further, the models that have been developed have never been adequately compared with each other using a consistent data source. This section describes the basic alternative models and compares them directly using, in part, the data on standardized earnings differentials that can be generated from the previous section.

Conceptual Models

Two basic classes of models have been considered, "compensating differential" models and "labor demand" models. The compensating differential model, in its general application, assumes that individuals completely respond to differences among areas to trace out a long-run equilibrium of observed earnings differentials that reflects the underlying attractiveness of areas or taste differences among workers. Consider, for example, a difference in price levels across areas. If price level is the only difference among areas and if workers accurately perceive and respond to such differences, an equilibrium would have real wages constant across areas; in other words, nominal wages would vary directly with price level. Similarly, from a worker's viewpoint, if the probability of unemployment varies across areas, expected real wages should vary inversely with employment probabilities. And, in general, if other area amenities vary, wages should vary in accordance with workers' valuations of these attributes. Taken together, for workers with identical tastes, wages should be related such that, for area j,

$$(5) \quad w^j = w^* + \alpha_1 p^j + \alpha_2 e^j + Z^j\gamma \qquad \text{with } \alpha_1 = 1 \text{ and } \alpha_2 = -1$$

where (in logarithms) w^j = nominal wage rate; w^* = real wage rate (a constant); e^j = employment rate; and Z^j = vector of other area attributes.

Past studies have generally concentrated on one or another of these factors. Commonly, a price coefficient of +1 is imposed (e.g., Fields [10] or Hall [17]).[20] Harris and Todaro [25] and Hall [17] consider the effect of different employment probabilities,[21] while Rosen [39] considers the more

20 Rosen [39] emphasizes estimates based upon a price coefficient constrained to +1, although his unconstrained estimates indicate an elasticity of +3 on prices.

21 Hall [17] actually estimates a linear model with unemployment rates instead of the logarithmic one in employment rates. This model, in terms of "risk premiums," neglects any effect of slack demand (indicated by unemployment rates in a cross-section) on wages and any effects of unemployment insurance.

general model that includes other area attributes such as climate, crime, and environment. Two important caveats, sometimes mentioned in the development of these models but often neglected, are that firms may also respond to area differences (cf. Hall [17] and Rosen [37]) and that workers may differ in preferences (cf. Rosen [39]); in either event, the hypotheses about the price and employment coefficients will differ from +1 and −1. Additionally, in the more general amenity model, there is little guidance in the definition and measurement of the relevant amenities.

The alternative models concentrate upon the characteristics or mix of employing industries in an area and (implicitly) assume impediments to geographic adjustments by firms and workers. One line of investigation considers how area characteristics affect production efficiency. The basic argument is that agglomeration economies—arising from availability of complementary businesses, more extensive labor markets, the development of social overhead capital, etc.—affect production efficiency in local areas. Agglomeration economies, proxied by both population (e.g., Sveikauskas [45], Segal [41]) and density (Hoch [27], Mera [31]) are then related to labor productivity in analyses of production functions.[22] (Analyses of income distribution and city differences (e.g., Garofalo and Fogarty [14]) frequently take a similar perspective with respect to productivity differences.)[23] An alternative line, albeit less well developed, considers directly how different industrial mixes of areas affect labor demands and wages (e.g., Scully [40]). Both types of investigations suggest that "quasi-rents" to workers in given areas, owing to and bounded by adjustment costs and locational fixities, are reflected in differential wages—even though they may disappear in the long run after complete adjustment by firms and workers.

22 Population and population density have been interpreted differently by some authors. Nordhaus and Tobin [36] and Kelley [29] interpret these as "total amenities" of an area, while Hirsch [26] suggests a "national distribution" model where population shifts local earnings distributions (although there is no theoretical justification for such shifts). Fuchs [11] finds that city size has a significant effect on hourly earnings and suggests that it may reflect either cost of living, labor quality, or disequilibrium conditions in local labor markets. Finally, related work on income distribution (e.g., Garofalo and Fogarty [14] or Hoch [27]) suggests both productivity and amenity interpretations of population and density.

23 Analyses of income distributions across cities (e.g., Garofalo and Fogarty [14] and the references there) are closely related. They suggest that population reflects above agglomeration economies and area amenities and estimate a nonlinear reduced form. In this, they also follow the past practice of "gross standardizatin" for schooling and other worker characteristics; see also Segal [41]. A final set of demand studies considers aggregate production functions and the substitution between schooling classes (see Section I and particularly Dougherty [6] and Johnson [28]). These ignore other area attributes and seem less appropriate for cross-sectional observations within a single open economy, although Johnson does jointly consider migration with wage determination.

TABLE 4
MODELS OF SMSA EARNINGS DIFFERENCES

	$\ln PV_{HS}$			$\ln PV_{COL}$		
Variable	(1)	(2)	(3)	(1)	(2)	(3)
ln *COST*	.541		.249	.443		.158
	(4.7)[a]		(6.3)[a]	(5.2)[a]		(7.5)[a]
ln *EMPLOY RATE*	−.182		−.396	−.101		−1.93
	(2.2)[b]		(1.6)[b]	(2.1)[b]		(1.3)[b]
ln *CRIME*	.111		.101	.044		.011
	(6.1)		(4.9)	(2.2)		(0.8)
ln *PARTICULATE*	.034		−.001	.015		−.010
	(1.8)		(0.1)	(0.7)		(0.5)
ln *SUN DAYS*	−.123		−.049	−.091		−.007
	(2.3)		(2.1)	(1.6)		(0.1)
MANUFACTURE		.001	.002		.001	.002
		(1.1)	(2.1)		(1.2)	(1.8)
CONSTRUCTION		−.023	−.006		−.007	.008
		(3.7)	(1.0)		(1.1)	(1.3)
GOVERNMENT		−.002	.000		−.002	−.001
		(1.6)	(0.2)		(1.4)	(0.8)
ln *POPULATION*			.033			.036
			(3.6)			(3.9)
ln *DENSITY*			−.007			.020
			(0.7)			(2.3)
Constant	6.38	12.14	8.60	7.97	12.17	9.97
R^2	.49	.35	.60	.37	.35	.58

Notes: t-statistics for null hypothesis H_o: $\beta=0$ except (a) H_o :$\beta=1$, and (b) H_o:$\beta=-1$.
Variable Definitions and Sources: PV = expected present value of earnings for high school and college graduates; *COST* = BLS Intermediate Budget estimate for SMSA; *EMPLOY RATE* = employment rate for male labor force members with ≤ 12 years of schooling (for *HS* models) or > 12 years of schooling (for *COL* models); *CRIME* = total crime rate (Liu [30]); *PARTICULATE* = rate of suspended particulates in air (Liu [30]); *SUN DAYS* = average number of sun days in year (Liu [30]); *MANUFACTURE* = percent total employment in manufacturing 1970; *CONSTRUCTION* = percent total employment in construction 1970; *GOVERNMENT* = percent total employment in all levels of government 1970; *POPULATION* = 1970 SMSA population; *DENSITY* = 1970 population density (persons/sq.mile). (Employment distributions and population variables from *County and City Data Book, 1972.*)

The compensating differential and labor productivity, or demand, models lead to very different interpretations of observed wage variations and are based upon quite different premises. Importantly, even within the two classes of models, there has been little research that consistently tests the alternative hypothesis; there has been no overlap between the classes.

This limited testing of the models is partially attributable to inadequate data but, probably to a greater extent, arises from the conceptual incompleteness of the models. The next section directly compares the alternatives using a consistent data set. The final section suggests how the models might be integrated.

Effects of Labor Market Attributes

Previous testing of these models has either ignored differences in characteristics of workers or used the "gross standardization" procedure depicted in equation (3). However, as demonstrated in Section II, variations in returns to individual worker characteristics cannot be neglected in considering the effects of local labor market factors. This analysis combines the information about the level and shape of earnings profiles by estimating the expected present value of lifetime earnings in each SMSA for workers with given schooling levels (here for high school graduates, or $S = 12$, and for college graduates, or $S = 16$).[24] Table 4 presents estimates of the various models of local earnings differences for the two schooling levels.

For the two schooling levels, three separate models are presented: a pure compensating differential model, a pure industrial mix model, and a combined model. In terms of explained variance, the pure compensating differential model (1) does better than the pure industrial mix model (2), but only marginally so for college graduates; however, the combined model (3) does significantly better than the separate ones. Importantly, substantial variation is unexplained even in the combined model.

The estimated models provide some interesting insights. In the pure compensating differential model (1), the estimated effects of price levels and employment probabilities are significantly different from the predictions in the simple models (i.e., from +1 and −1, respectively). These estimates imply significant adjustment lags for individuals and, probably, more complex labor market adjustments that include firm behavior.[25]

---

24 Present value of lifetime earnings, while perhaps not appropriate for individual decisions because of lifetime migration patterns, is used to summarize the different dimensions of local earnings patterns. The estimation pertains to the 141 SMSAs for which complete data were available. While only OLS estimates are presented, GLS estimates that allowed for the differential sampling variances in the estimation of present values (cf. Hanushek [22]) were also obtained but were qualitatively similar to the OLS estimates.

25 If only the employment rate is included in the equation, the estimated coefficients are insignificantly different from −1. This finding reinforces a major criticism of past analyses: Taken individually, each hypothesis receives much more support than when considered in a more general framework. This appears to be more than simply data problems. Even by itself, the price term is always significantly less than +1. This may partly be caused by measurement errors; however, estimates based upon the U.S. Chamber of Commerce cost-of-living index (Liu [30]) yielded very similar results. The

There is no consensus on other area attributes that should be included, and the specification of these (crime rates, suspended particulates in the air, sunny days) follows that of Rosen [39]. The effects of these amenities are as expected: wages are lower in more pleasant and safer places. However, while there are no real expectations about the magnitudes of these coefficients, some caution in their interpretation is suggested by their being smaller in magnitude for the more mobile college workers.[26]

The pure industrial mix models, on the other hand, do surprisingly well at explaining the same earnings differences. With no consideration of anything except three aggregate employment distribution measures, 35 percent of the variation is explained; for college-level workers, this is virtually the same as the pure compensating differential models.

The most interesting models, however, are the combined ones (3). These indicate that both explanations of earnings variation appear important, even if the precise interpretation is cloudy. The estimates for the compensating differentials, while significant, remain different from those hypothesized: Price and employment rate coefficients are small and significantly different from those suggested by individual equilibrium, and the effect of other specific amenities are uniformly larger for the less mobile class of workers. Industrial composition, specifically manufacturing employment, indpendently raises wages. However, the most important difference in wages appears to be related simply to SMSA population, and possibly density.[27] Here is where the serious interpretation problems enter.

Population or density was discussed as affecting productivity and labor demand, but others have simultaneously used these as proxies for area amenities (see fn. 22). As the models have been formulated in the past, there is no way to distinguish between these two different interpretations—even though the two differ importantly in assessing welfare implications, in

---

BLS budget data are available for only 34 of the included SMSAs; for the remainder, the regional metropolitan area mean was used (except that Honolulu was excluded from the western states). When tested, there was an insignificant difference in price coefficients between "known" and "estimated" prices. Finally, estimates also included the low-budget estimate in place of and in addition to the intermediate-budget estimates presented. In all cases, the separate and combined estimates are significantly below +1.

26 Even if the underlying model were correct, this could result from taste differences (cf. Rosen [39]) or from measurement problems. No measure of location or amenity differences within SMSAs are included; since college-educated workers tend to live in suburban areas more frequently, SMSA measures, say, of crime rates, may not reflect relevant amenities for them. Nevertheless, since estimates are based upon within-schooling-group data, SMSA rates may still be a reasonable index across areas of relevant amenities.

27 Approximately half of the earnings variance can be explained by the single variable—SMSA population.

evaluating various public policies, and in analyzing such behavior as individual and firm migration.

The basic point is rather straightforward: While there appear to be significant and stable differences in earnings relationships across labor markets, there are real questions about how to explain them. Existing data clearly make testing alternative theories difficult; even with the more extensive and refined data used here, intercorrelations among labor market attributes affect tests of the alternative models—ones based upon quite different perspectives and incompatible assumptions. However, a more serious problem is the inadequacy of existing specifications. The models considered are really reduced-form models, of the form suggested in equation (2). In these, it is difficult to introduce restrictions that would allow distinguishing among alternative explanations. Past analyses, focusing on one or another of the regional hypotheses, appear to confirm each when analyzed in isolation, but, when jointly considered, this confirmation evaporates. Tests of existing models, including those presented here, are simply not very powerful and do not give much information about the validity of the underlying hypotheses. This is not so much a statistical and data problem as it is a problem in modeling the underlying behavior that generates the observed earnings differences.

IV. OUTLINE FOR FUTURE RESEARCH

The key to understanding individual earnings determination lies in the specification of the structural relationships which underlie equation (2). These relationships indicate how differences in the supply and demand for labor interact to determine the rewards to individual characteristics. Without geographically detailed individual data, such investigations have not been feasible. However, with the increased availability of such data, a more structural approach now seems warranted and productive.

Earnings analyses have highlighted the importance of heterogeneity of the labor force and the differential earnings of individuals with varying skills and training. The supply-side analysis should relate to the different compositions of local labor forces, while the demand-side analysis must consider the different substitution possibilities across skill classes.

There are two components of the supply side, migration and schooling decisions. Most migration research (Greenwood [15]) has taken a disequilibrium view; that is, that migration is a response to differences in earnings opportunities across areas. However, most of this has looked at aggregate moving (across all labor classes) and has neglected the local differences in earnings by schooling and age noted in Section II. There is also the possibility that a portion of migration represents an "equilibrium" response to amenity differences.[28] The effect of local earnings differences on schooling decisions has received little attention.[29] Both of these can, with currently available data, be considered in much more detail than has previously been done.

The demand side logically begins with production relationships and the substitutability of different skill groups. Past research into manufacturing industries has suggested different substitution relationships among industries, but direct analyses of labor demand have assumed a common aggregate demand function. Because of industry differences in production functions and different local industrial compositions, the aggregate demand for classes of labor within a local area will differ and, given different relative supplies of labor, should lead to different wage structures across areas (as found in Section II).[30] If a common production function for each industry is assumed, it is possible to aggregate these for each area on the basis of local industrial compositions and thus to specify the aggregate demand relationships within areas.

Important parts of research in each of these areas have been analyzed. Yet they have not been put together in a consistent and coherent manner in the analysis of structural aspects of labor market operations. Simultaneously analyzing all of the aspects is a truly formidable task. However, the more modest goal of recognizing the common structure to the separate analytical efforts and developing the individual pieces within this structure is currently feasible. In part, this simply calls for capitalizing upon the greatly expanded micro-data sets now available.

---

28 For example, there may be life-cycle migration patterns that reflect changing evaluations of income and other area attributes. In such a case, a more general definition of "income" is required, and more attention must be given to life-cycle characteristics of migrants.

29 Consistent with past earnings analysis, most research into schooling decisions has neglected any earnings differences across areas. However, at the very least, schooling decisions must recognize the varying opportunity costs of schooling across areas. Some recent analyses of college decisions (Dresch, Hanushek, and Waldenberg [8]) show that earnings differences do lead to cross-sectional variations in college attendance.

30 Derived demand analyses (see fn. 4) commonly assume an aggregate CES production function. With this and exogenously given labor supplies, relative wages by labor class are directly related to relative supplies with the relationship being given by the elasticity of substitution. This has also been extended by Johnson [28] to consider simultaneously determined migration. When this approach, which imposes strong restrictions on local demand, is followed using the SMSA data above, implausible estimates of substitution elasticities are obtained, and these estimates are very sensitive to the precise specification of the model. This analysis is also inconsistent with the direct estimation of manufacturing production functions; the latter assumes relative wage differences across areas. Estimates of elasticities of substitution based upon simple demand models range from .85 to 43.5, depending upon specification and estimation method, for the SMSAs used in Section III. (These are available from the author.)

REFERENCES

1. Mark Blaug. "The Empirical Status of Human Capital Theory: A Slightly Jaundiced Survey." *Journal of Economic Literature* 14 (September 1976): 827–55.
2. Joseph W. Bloch. "Regional Wage Differentials: 1907–1946." *Monthly Labor Review* 60 (April 1948).
3. George H. Borts. "The Equalization of Returns and Regional Economic Growth." *American Economic Review* 50 (June 1960): 319–47.
4. Samuel Bowles. "Aggregation of Labor Inputs in the Economics of Growth and Planning: Experiments with a Two-Level CES Function." *Journal of Political Economy* 78 (January/February 1970): 68–81.
5. Barry R. Chiswick. *Income Inequality.* New York: National Bureau of Economic Research, 1974.
6. C. R. S. Dougherty. "Estimates of Labor Aggregation Functions." *Journal of Political Economy* 80 (November/December 1972): 1101–19.
7. Stephen Dresch. "Demography, Technology, and Higher Education." *Journal of Political Economy* 83 (June 1975): 535–69.
8. Stephen Dresch, Eric Hanushek, and Adair Waldenberg. "Economic Incentives and College Entry." New Haven, CT: Institute for Demographic and Economic Studies, 1979, mimeo.
9. P. R. Fallon and Richard Layard. "Capital-Skill Complementarity, Income Distribution, and Growth Accounting." *Journal of Political Economy* 83 (April 1975): 279–302.
10. Gary S. Fields. "Labor Force Migration, Unemployment and Job Turnover." *Review of Economics and Statistics* 42 (August 1960): 292–300.
11. Victor Fuchs. *Differentials in Hourly Earnings by Region and City Size.* Occasional Paper 101. New York: National Bureau of Economic Research, 1967.
12. Victor Fuchs and Richard Perlman. "Recent Trends in Southern Wage Differentials." *Review of Economics and Statistics* 42 (August 1960): 292–300.
13. Lowell E. Gallaway. "The North-South Wage Differential." *Review of Economics and Statistics* 45 (August 1963): 264–72.
14. Gasper Garofalo and Michael S. Fogarty. "Urban Income Distribution and the Urban Hierarchy-Equality Hypothesis." *Review of Economics and Statistics* 61 (August 1979): 381–88.
15. Michael J. Greenwood. "Research on Internal Migration in the United States: A Survey." *Journal of Economic Literature* 13 (June 1975): 392–433.
16. Zvi Griliches. "Notes on the Role of Education in Production Functions and Growth Accounting." In *Education, Income, and Human Capital,* ed. W. Lee Hansen. New York: National Bureau of Economic Research, 1970.
17. Robert E. Hall. "Turnover in the Labor Force." *Brookings Papers on Economic Activity* (3: 1972).
18. Frank A. Hanna. "Contribution of Manufacturing Wages to Regional Differences in Per Capita Income." *Review of Economics and Statistics* 33 (February 1951): 18–28.

19. ———. *State Income Differentials, 1919–1954.* Durham, NC: Duke University Press, 1959.
20. Giora Hanoch. "An Economic Analysis of Earnings and Schooling." *Journal of Human Resources* 2 (Summer 1967): 310–29.
21. Eric A. Hanushek. "Regional Differences in the Structure of Earnings." *Review of Economics and Statistics* 55 (May 1973): 204–13.
22. ———. "Efficient Estimators for Regressing Regression Coefficients." *American Statistician* (May 1974).
23. ———. "Ethnic Income Variations: Magnitudes and Explanations." In *American Ethnic Groups,* ed. Thomas Sowell. Washington: Urban Institute, 1978.
24. Eric A. Hanushek and John M. Quigley. "Implicit Investment Profiles and Intertemporal Adjustments of Relative Wages." *American Economic Review* 68 (March 1978): 67–79.
25. John R. Harris and Michael P. Todaro. "Migration, Unemployment, and Development: A Two Sector Analysis." *American Economic Review* 60 (March 1970): 126–42.
26. Barry T. Hirsch. "Predicting Earnings Distributions Across Cities: The Human Capital Model vs. the National Distribution Hypothesis." *Journal of Human Resources* 13 (Summer 1978): 366–84.
27. Irving Hoch. "Income and City Size." *Urban Studies* 9 (October 1972): 299–328.
28. George E. Johnson. "The Demand for Labor by Educational Category." *Southern Economic Journal* 37 (October 1970): 190–204.
29. Kevin C. Kelley. "Urban Disamenities and the Measure of Economic Welfare." *Journal of Urban Economics* 4 (October 1977): 379–88.
30. Ben-Chieh Liu. *Quality of Life Indicators in U.S. Metropolitan Areas, 1970.* Washington: U.S. Government Printing Office, 1975.
31. Koichi Mera. "On the Urban Agglomeration and Economic Efficiency." *Economic Development and Cultural Change* 21 (January 1973): 309–24.
32. Jacob Mincer. "The Distribution of Labor Incomes: A Survey." *Journal of Economic Literature* 8 (March 1970): 1–29.
33. ———. *Schooling, Experience, and Earnings.* New York: National Bureau of Economic Research, 1974.
34. National Bureau of Economic Research. *Regional Income.* Studies in Income and Wealth, Vol. 21. Princeton, NJ: Princeton University Press, 1957.
35. Richard R. Nelson and Edmond Phelps. "Investment in Humans, Technological Diffusion and Economic Growth." *American Economic Review* 56 (May 1966): 69–75.
36. William Nordhaus and James Tobin. "Economic Growth." In *National Bureau of Economic Research Fiftieth Anniversary Colloquium,* Vol. V. New York: Columbia University Press, 1972.
37. Sherwin Rosen. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 82 (January 1974): 34–55.

38. ———. "Human Capital: A Survey of Empirical Research." In *Research in Labor Economics,* ed. Ronald G. Ehrenberg. Greenwich, CT: JAI Press, 1977.
39. ———. "Wage-Based Indexes of Urban Quality of Life." In *Current Issues in Urban Economics,* eds. Peter Mieszkowski and Mahlon Straszheim. Baltimore: Johns Hopkins University Press, 1979.
40. Gerald W. Scully. "Interstate Wage Differentials: A Cross Section Analysis." *American Economic Review* 59 (December 1969): 757–73.
41. David Segal. "Are There Returns to Scale in City Size?" *Review of Economics and Statistics* 58 (August 1976): 339–350.
42. Martin Segal. "Regional Wage Differences in Manufacturing in the Postwar Period." *Review of Economics and Statistics* 43 (May 1961): 148–55.
43. James P. Smith and Finis Welch. "The New Depression in Higher Education: Ulcers or Indigestion." Mimeo, 1977.
44. Mahlon Straszheim. "An Introduction and Overview of Regional Money Capital Markets." In *Essays in Regional Economics,* eds. John F. Kain and John R. Meyer. Cambridge, MA: Harvard University Press, 1971.
45. Leo Sveikauskas. "The Productivity of Cities." *Quarterly Journal of Economics* 89 (August 1975): 393–413.
46. Lester Thurow. *Poverty and Discrimination.* Washington: Brookings Institution, 1969.
47. T. Nicholas Tideman. "Defining Area Stress in Unemployment." *Public Policy* 21 (Fall 1973): 441–92.
48. Finis Welch. "Measurement of the Quality of Schooling." *Amerian Economic Review* 56 (May 1966): 379–92.
49. ———. "Education in Production." *Journal of Political Economy* 78 (January/February 1970): 35–59.
50. ———. "Black-White Differences in Returns to Schooling." *American Economic Review* 63 (December 1973): 893–907.