Intra-urban Earnings Differences: Spatial Mismatch or Selective Migration?

ENGBERG, JOHN AND KIM, TAEIL
The H. John Heinz School of Public Policy and Management
Carnegie Mellon University

Abstract

We pose two models for earnings differences between residential locations in an urban area. The spatial mismatch model is based on barriers to labor mobility and suggests that observed earnings differences are real. The selective migration model suggests that observed earnings differences reflect unobserved differences in individual abilities. Using a sample of white men with a high school degree or less, we estimate a model of location choice and location-specific earnings for three portions of the Allegheny County (Pittsburgh). Our results indicate that the earnings decline in the old manufacturing portion of the county reflects selective emigration rather than continuing excess labor supply.

I. Introduction

Employment in the U.S. has undergone a long shift from manufacturing to service industries (deindustrialization) and from central city to suburban locations (job suburbanization). While this economic restructuring process has resulted in economic growth for the suburbs, it has hastened the deterioration of central cities and manufacturing communities, especially in the old manufacturing region. In metropolitan Pittsburgh, the case we analyze, the decay is obvious in the towns surrounding the closed steel mills and in the city's low income neighborhoods.

The relocation of economic activity has, in many cases, been mirrored by changes in

The Korean Journal of Policy Studies Vol. 11 (1996), 105~127

residential patterns. As jobs move to the suburbs, many people follow, either through commuting or migration. Labor supply follows the demand, with wages adjusting in response. If mobility is costless, we would expect wages for people with comparable skills to be fairly equal among portions of a metropolitan area. However, in Pittsburgh, this does not appear to be the case. Between 1980 and 1990, average real earnings for white men with a high school education or less has declined 22% if they live in the old manufacturing communities, but only declined 12% in the expanding suburbs.

We examine two possible explanations for intra-metropolitan earnings differences. The first explanation borrows from the spatial mismatch literature. If there are barriers to worker mobility, differences in labor demand between locations will yield real earnings differences. The second explanation focuses on selective migration. If the most productive workers choose to reside in the expanding suburbs, average earnings will be higher in the suburbs.

The spatial mismatch hypothesis (Kain, 1968) states that apparent labor market discrimination suffered by blacks is due instead to housing market barriers. As jobs relocated to the suburbs, housing segregation prevented blacks from following, leading to lower employment and lower wages. The empirical evidence for the spatial mismatch hypothesis is mixed, but the hypothesis remains a prominent theory of spatial differences in labor market outcomes.[1)]

Our interest here is not in housing segregation, but in other barriers to labor mobility. The arbitrage of wage differences by worker mobility requires that individuals have information about the spatial distribution of demand differences and that migration and/or commuting costs are minimal. The cost of information and moving may be negligible for individuals with abundant human capital but be relatively large for low skill individuals. This implies that earnings differences due to spatial mismatch may exist for low skilled whites. Although there are many spatial mismatch studies which focus on earnings of blacks, we know of no spatial mismatch research on whites.[2)]

On the other hand, spatial differences in earnings could reflect real differences in average abilities. In the absence of barriers to labor mobility, wages will not differ by residential location for people of equal productivity. However, standard proxies for human capital, such as education level and age, only capture a limited portion of differences in individual productivity. Among people with the same observable characteristics, more productive individuals might choose to live in the expanding suburbs, perhaps due to a package of residential amenities with a positive income elasticity. In such a situation, earnings differences do not reflect mobility barriers but instead are the result of real differences in productivity.

In this paper, we investigate the spatial earnings differences for low skilled white males. We analyze samples from the 1980 and 1990 Public Use Micro Sample (PUMS) Census data for the Pittsburgh metropolitan area. To determine whether earnings differences can be attributed to mobility barriers, we estimate separate earnings regressions for three distinct portions of the metropolitan area. We refer to spatial differences in the earnings regressions as "place effects." However, we allow residential location to be endogenous. We use sample selection methods to capture unobserved differences in individual productivity which we term "sorting effects."

We study the Pittsburgh metropolitan area because it is representative of the old manufacturing areas which have undergone a long decline in manufacturing. Steel production, which was the major industry of Pittsburgh, has declined for several decades. The gradual shift of U.S. steel production from Pittsburgh to Chicago area began after World War I. After a temporary halt during the Second World War, the trend toward deindustrialization resumed after the war and accelerated in the 1970's and continued during 1980's. During the period we study, 1980 to 1990, employment in steel industry decreased 70% and total number of manufacturing jobs declined 44% in the Pittsburgh Metropolitan area. Over the same period, steel industry employment and the total number of manufacturing jobs in the U.S. fell by 50% and by 9%, respectively.3)

Pittsburgh also provides a unique opportunity to disentangle the issues of economic change and education from the issue of race. Compared with other "rust belt" metropolitan areas, racial minorities make up a very small portion of the population. In 1990, whites comprise over 90 percent of the population.4) Therefore, the Pittsburgh area is well suited to provide a baseline estimate of spatial effects due to the restructuring process of U.S. economy.

Our methods follow much the spatial mismatch literature in comparing the labor market outcomes between portions of a metropolitan area. However, our focus on selective migration as an alternative explanation and on distinct patterns for old manufacturing communities leads us to address the following two concerns.

First, with few exceptions,5) most previous spatial mismatch studies ignore the endogeneity of location choice. The place effects are estimated by comparing the earnings of central city residents with otherwise similar suburban residents. This will provide inconsistent estimates, however, if location is correlated with unobserved individual characteristics which have direct effects on earnings. Since an individual's economic status may affect his or her location choice, treating location choice as exogenous could result in serious biases when estimating the impact of the location on economic status. For example, while living in the suburbs may increase earnings because good jobs are more accessible, it is also likely that people with higher earnings are more inclined to select a suburban residence (Ihlanfeldt and Sjoquist 1990). If the location choice is not exogenous, estimation of the direct impact of location requires an estimation method which accounts for the endogeneity of location.

Second, the simple central city/suburban dichotomy is overly restrictive when analyzing intra-metropolitan variation (Ihlanfeldt and Sjoquist 1989). The traditional spatial mismatch approach which divides workers into central city and suburbs does not allow for differences in economic growth across suburbs. For example, in the Pittsburgh metropolitan area, the suburbs include not only expanding communities, but also the communities which contained the old steel mills. In this paper, we divide the suburbs into the old manufacturing area and the expanding suburbs. We compare the expanding suburbs to both the central city and to the old manufacturing area.

Briefly, our strategy is to examine the impact of allowing for endogenous location when estimating location-specific earnings regressions. Using 1980 and 1990 PUMS data for prime age (25-55) white male householders who work full time, we estimate two versions of the model. The first, which does not allow for the residential location to be endogenous, implies that high average earnings for residents of the expanding suburbs are almost entirely due to an unexplained wage premium -- i.e. place effects. The second, which allows for the endogeneity of residential location, implies that high suburban wages reflect the choice of productive workers to live in the suburbs. This suggests that 'spatial mismatch' is not a problem among low skill white male workers.

In the following section, the sample selection model which accommodates the endogeneity problem of location choice is discussed. Section III describes the Pittsburgh sample. Sections IV and V present bivariate analysis and regression results, respectively. A summary and discussion are provided in the last section.

II. Model

To begin with, assume that there are only two locations,[6] suburbs and central city. Each member of the population chooses one or the other. A pair of equations describe potential earnings in each location and a third equation models the location choice. The two parts of the model are linked in that potential earnings is one of the factors determining location choice.

The potential earnings of individual i, who is randomly drawn from a population, may differ in the two locations. If he resides in the suburbs,

$$W^{*}_{si} = X_i\beta_s + \varepsilon_{si} \tag{1}$$

and if he resides in the central city,

$$W^*_{ci} = X_i\beta_c + \varepsilon_{ci} \tag{2}$$

where W*s and W*c are log (potential) earnings for residents of the suburbs and in the central city, respectively. The vector Xi represents observable individual (productivity) characteristics, and the disturbance terms ε si and ε ci represent unobserved characteristics which are assumed to be normally distributed with zero means, variances σs2 and σc2 covariance σsc. Actual earnings, Wi, are only observed for the location in which an individual resides.

Potential earnings will differ if the coefficients differ between the locations. A larger coefficient in the suburbs than in the city for a particular element of Xi indicates a bigger return to that characteristic in the suburbs. For example, if the coefficient on education is larger in the suburbs, then the suburban resident's labor market values the productive abilities of highly educated workers more than the labor market for city residents. A larger intercept in the suburbs suggests that high wages in the suburbs are not related to particular skills.

In a similar fashion, the variances and covariance of ε si and ε ci provide information about the two labor markets. A higher variance for suburban residents indicates that a larger portion of the variation in wages is due to productive characteristics that are unobserved. A correlation close to unity suggests that the unobserved factors that are important in the suburbs are the same as those that are important in the city, while a low correlation indicates that the suburbs reward unobserved characteristics that are not valued in the city.

The residential location is determined by comparing the utility person i receives from living in the suburbs to that received from living in the central city:

$$U_{si} = W^*_{si}\gamma_s + A_{si} \tag{3}$$

$$U_{ci} = W^*_{ci}\gamma_c + A_{ci} \tag{4}$$

Usi and Uci are the utilities in the suburbs and in the central city, respectively. The coefficients γs and γc, assumed to be positive, are the marginal utilities of earnings in each location. The location with a relatively large value of γ is more highly valued by high income individuals, perhaps due to the level of public goods and associated taxes (Engberg and Kim, 1993). The terms Asi and Aci represent the utilities from other amenities of the residential locations.

We assume that the differences in the utilities of amenities (Asi - Aci) can be decomposed into a portion that depends on observed individual characteristics and an unobserved component:

$$A_{si} - A_{ci} = Z_i\alpha + v_i \tag{5}$$

The vector Zi contains observable characteristics, including the characteristics in the earnings equation, Xi. The error term, vi, represents unobserved factors which affect valuation of amenities in the two locations. It is assumed to be normally distributed with mean zero, standard deviation σv and covariances σsv and σsv with ε si and ε ci.

Let the difference in utility obtainable in the city and suburbs be Ui*=Usi - Uci. We assume a person i chooses to live in the suburbs if Ui* > 0 and chooses the central city otherwise.7)

Repeated substitution gives a reduced-form location choice equation:

$$\begin{aligned} U_i^* &= (W_{si}^*\gamma_s - W_{ci}^*\gamma_{c}) + Z_i\alpha + v_i \\ &= Q_i\theta + u_i \end{aligned} \tag{6}$$

The vector Qi contains all the unique elements of Xi and Zi. The parameter vector θ captures the indirect effect through wages of Xi as well as the direct effect of Xi and Zi on location choice. The composite disturbance term contains the disturbances in the earnings and location choice equations: ui= ε si γ s- ε ci γ c+vi. The assumptions on the individual equations imply that the composite disturbance term is normally distributed; therefore, equation (7) can be estimated by maximum likelihood probit methods. The location choice coefficients can only be identified up to scale, so, as is customary, we normalize by setting the variance of ui to unity.

We have assumed that, for the entire population, unobserved factors which affect wages,

ε, are uncorrelated with observed factors, X. However, within any location, ε and X are likely to be correlated. The choice process implies that the subpopulation of individuals who choose a particular location will have correlated values of X and ε. For example, if an individual who moves to an affluent area has a low value of observable income-producing traits (X), he is likely to have a high value of unobservable income-producing traits (ε). This within-area negative correlation of X and ε implies that ordinary least squares regression does not yield consistent estimates of the area-specific wage equations (1) and (2). Therefore, the estimates of the location choice equation are used to create selection-correction terms for the earnings equations (Heckman 1976, Lee 1978). Including selection correction terms, the earnings equations become:

$$W_{si} = X_i\beta_s + \sigma_{su}\lambda_{si} + \eta_{si} \tag{7}$$

$$W_{ci} = X_i\beta_c + \sigma_{cu}\lambda_{ci} + \eta_{ci}] \tag{8}$$

where

$$\lambda_{si} = \frac{f(Q_i\theta)}{F(Q_i\theta)} = E(u_i | U^* > 0) > 0, \text{ and } \lambda_{ci} = \frac{-f(Q_i\theta)}{1 - F(Q_i\theta)} = E(u_i | U^* \leq 0) < 0 \tag{9}$$

The selection correction terms, λsi and λci, are related to the location choice equation. A large value of either selection term indicates that the person is likely to have unobservable characteristics, ui, that make him prefer the suburbs. The coefficients, σsu and σcu, on the selection terms measure the covariance between ui and the unobservable factors in the two wage equations, εsi and εci. By construction, the new error terms, ηsi and ηci, are uncorrelated with the explanatory variables within the subgroup choosing each location. This permits consistent estimation of the coefficients in the equations.

This model nests many explanations of spatial earnings differences, including the spatial mismatch and selective migration models presented in the introduction. A spatial mismatch model suggests that increased labor demand in a particular area is reflected by higher wages in the area. Therefore, if the suburbs have higher labor demand, the coefficients of the suburban residents' earnings equation should imply higher expected wages than the coefficients of the city residents' earnings equation.

On the other hand, increased labor demand in an area could lead to an inflow of workers. If mobility is sufficient to eliminate excess demand, the coefficients on observable and unobservable characteristics in the earnings equations will be the same for the two areas.

However, high labor demand in the suburbs will still lead to high average wages if individuals with higher values of observable and unobservable productive attributes choose to reside in the suburbs. In this case, the coefficients, σsu and σcu, on the selection terms will be positive in both equations because the unobservable characteristics which make an individual likely to live in the suburbs are productive attributes and are (positively) rewarded in both areas.

As mentioned above, this type of sample selection is different from the frequently cited

Roy Model (1951) in which the choice of labor market is determined by comparative advantage. If the labor markets for suburban and city residents value different skills, including different unobservable individual attributes, we will observe a positive coefficient on the selection term in the suburban equation and a negative coefficient on the selection term in the city equation.

The selection correction terms have important implications for measuring the magnitude of place effects. The location-specific expected earnings of an individual with attributes Xi can be determined in both the chosen location and the alternative location:

$$E(W_s|X_i, U^* > 0) = X_i\beta_s + \sigma_{su}\lambda_{si} \tag{10}$$

$$E(W_s|X_i, U^* \leq 0) = X_i\beta_s + \sigma_{su}\lambda_{ci} \tag{11}$$

$$E(W_c|X_i, U^* > 0) = X_i\beta_c + \sigma_{cu}\lambda_{si} \tag{12}$$

$$E(W_c|X_i, U^* \leq 0) = X_i\beta_c + \sigma_{cu}\lambda_{ci} \tag{13}$$

Note that actual earnings are only observed for equations (10) and (13).

The average earnings differences between residents of the suburbs and central city can be decomposed in the following manner:

$$\begin{aligned} \overline{W}_s - \overline{W}_c &= E(W_s|\ \overline{X}_s, U^* > 0) - E(W_c|\ \overline{X}_c, U^* \leq 0) \\ &= [E(W_s|\ \overline{X}_c, U^* \leq 0) - E(W_c|\ \overline{X}_c, U^* \leq 0)] + [E(W_s|\ \overline{X}_s, U^* > 0) - E(W_s|\ \overline{X}_c, U^* \leq 0)] \\ &= [\ \overline{X}_c(\beta_s - \beta_c) + (\sigma_{su} - \sigma_{cu})\ \overline{\lambda}_c] + [\beta_s(\ \overline{X}_s - \ \overline{X}_c) + \sigma_{su}(\ \overline{\lambda}_s - \ \overline{\lambda}_c)], \end{aligned} \tag{14}$$

where $\overline{X}_j$ are the average attributes in location j. The term $\overline{\lambda}_j$ is the average selection correction term which is equal to the average value of ui for people in location j.

The first set of bracketed terms in the second and third lines,

$$E(W_s|\ \overline{X}_c, U^* \leq 0) - E(W_c|\ \overline{X}_c, U^* \leq 0) = \ \overline{X}_c(\beta_s - \beta_c) + (\sigma_{su} - \sigma_{cu})\ \overline{\lambda}_s$$

provides the hypothetical wage gain to the average central city resident of moving from the central city to the suburbs. This wage gain comes from the differences in returns to observed characteristics (βs-βc) and the differences in returns to unobserved characteristics (σsu-σcu). Therefore, it can be defined as the place effect. This term would be positive in a spatial mismatch model and zero in a selective migration model.

The second set of bracketed terms,

$$E(W_s|\ \overline{X}_{s,} U^*>0)-E(W_s|\ \overline{X}_{c,} U^*\leq 0)=\beta_s(\ \overline{X}_s-\ \overline{X}_c)+\sigma_{su}(\ \overline{\lambda}_s-\ \overline{\lambda}_c),$$

indicates the extent to which the differences in average observed characteristics ($\overline{X}_s-\ \overline{X}_c$)and average unobserved characteristics ($\overline{\lambda}_s-\ \overline{\lambda}_c$) are rewarded in the suburbs. This is the component of average wage differences that can be attributed to location choice. The sum of these terms would be positive in the selective migration model.

Previous studies which ignore the endogeneity of location choice implicitly set the selection correction coefficient to zero. Such a procedure will attribute all wage differences to (inconsistently estimated) differences in returns to observable characteristics and to average differences in observable characteristics:

$$E(W_s|\ \overline{X}_{s,} U^*>0)-E(W_s|\ \overline{X}_{c,} U^*\leq 0)=(\beta_s^{OLS}-\beta_c^{OLS})\ \overline{X}_c+\beta_s^{OLS}(\ \overline{X}_s-\ \overline{X}_c), \tag{15}$$

where βsOLS and βcOLScare the ordinary least squares estimates obtained from suburban residents and central city residents, respectively. If there is no correlation between the wage equation errors and the location choice equation error (i.e.σsu=σ cu=0), then equations (14) and (15) are identical. In such a case, the OLS estimates of the coefficients are consistent and the place effects and sorting effects can be correctly estimated. However, if correlation exists, then OLS estimates would not be consistent and, therefore, the place effects and sorting effects are not correctly measured.

III. Data

The data for this study come from the 5% Public Use Micro Sample (PUMS) of 1980 and 1990 for Allegheny County in the Pittsburgh Metropolitan area. These data contain individual records from the 1980 and 1990 Censuses of Population and Housing. Allegheny County is the main county of Pittsburgh Metropolitan area. It includes City of Pittsburgh (central city), and Monogahela Valley, the former site of most of the area's steel production, as well as newer suburbs. About 70% of Pittsburgh metropolitan area residents live in Allegheny County.

From the data we select prime age (25-55) white male householders who worked full time (at least 35 hours per week) and year round (at least 40 weeks) and earned at least $100 per week in 1979 and in 1989, respectively. Also, to focus on low skilled workers, we restrict our sample to men who have no education past to high school graduation (12 years).[8] These restrictions yield a fairly homogenous group of workers.

In the PUMS data, a county is divided into sub-county areas which are named County Group Residential Places (CGRP) in 1980 data and Public Use Microdata Areas (PUMA) in 1990 data. We group these sub-county areas into three locations: the central city, the old (steel) manufacturing area and other suburban areas.[9] The following section describes the differences in earnings and other characteristics of residents of the three areas in 1980 and in 1990.

IV. Descriptive Statistics

<Table 1A> gives the means and standard deviations of the variables in the sample. For each variable, the first row is for the 1980 sample and the second row is for the 1990 sample. The wage data are given in 1989 dollars.[10] The average wage decreased about 15% in real terms from 1979 to 1989. The wage drop is particularly large in the old manufacturing area (22%) and in the central city (17%). As a result, the wage gap between the suburbs and the central city increased from 15% to 23% and the wage gap between the suburbs and the old manufacturing area increased from 3% to 17%. Table 1B compares the average wage between 1980 and 1990 by education level. The average wage of highly educated workers increased about 5% in Allegheny County while the average wage of poorly educated workers decreased about 15%. This is consistent with other recent studies which have found that the earnings gaps between poorly educated and well educated workers has increased over the 1980s (e.g. Acs and Danziger, 1990, Bound and Freeman, 1992, Card and Krueger, 1992, and Katz and Murphy, 1992).

<Table 2A> provides the change in industry of employment, by residential location, between
1980 and 1990. In all three areas, the manufacturing sector declined, while the trade and service sector grew. As expected, the largest decrease of employment in the manufacturing sector occurred in the old manufacturing area. The central city had the largest increase in service industry employment. Table 2B gives wage information by industry and residential location. In the central city and in the old manufacturing area, the average wage of the manufacturing sector is the highest among the four industry categories. The wage gaps between the locations are smallest in the manufacturing sector.

In summary, the above tables show that the Allegheny County labor market for low educated workers deteriorated between 1980 and 1990. Of the three areas, the deterioration was the most severe in the old manufacturing communities. The wage deterioration of poorly educated workers was associated with a shift in employment from manufacturing to other industries.

⟨Table 1A⟩ Prime Age Male Full Time Workers in Allegheny County, Pennsylvania Average Characteristics (Standard Deviations in Parentheses) ACD1980 (Top Entry) and 1990 (Bottom Entry)

	Location			
	Central City	Old Manufact	Suburbs	Total
Wage ($)	30145 (12355) 24997 (10983)	33748 (10794) 26285 (11195)	34746 (13139) 30716 (17719)	33539 (12497) 28516 (15350)
Weeks of work	50.8 (2.77) 51.1 (2.65)	51.0 (2.54) 51.1 (2.40)	51.1 (2.40) 51.2 (2.33)	51.0 (2.52) 51.2 (2.41)
Hours of work	41.9 (4.89) 42.7 (6.03)	42.4 (6.58) 43.6 (7.28)	42.7 (6.33) 43.6 (7.19)	42.5 (6.14) 43.4 (7.02)
Schooling	11.1 (1.47) 11.7 (0.91)	11.5 (1.25) 11.8 (0.81)	11.4 (1.30) 11.8 (0.77)	11.4 (1.33) 11.8 (0.81)
Age	41.0 (9.33) 39.7 (8.93)	41.5 (9.14) 40.1 (8.86)	41.6 (9.10) 41.3 (8.51)	41.4 (9.16) 40.7 (8.70)
Marriage	0.87 (0.33) 0.80 (0.40)	0.89 (0.31) 0.84 (0.37)	0.91 (0.29) 0.85 (0.36)	0.90 (0.31) 0.84 (0.37)
Sample Size	893 393	1212 558	2299 1195	4404 2146

Wage : yearly wage (real term in 1989)
Weeks of work : weeks of worked per year
Hours of work : hours of worked per week
Schooling : years of education
Marriage : =1 if spouse present, =0 otherwise

〈**Table 1B**〉 Average Earnings by Education Level 1980 (Top Entry) and 1990 (Bottom Entry)

	Location			
	Central City	Old Manufact	Suburbs	Total
High School Degree or less	30145 (12355) 24997 (10983)	33748 (10794) 26285 (11195)	34746 (13139) 30716 (17719)	33539 (12497) 28516 (15350)
More than High School Degree	38230 (22989) 39418 (32867)	40947 (19792) 37161 (23659)	46617 (23602) 50469 (38912)	44051 (23011) 46117 (35945)

〈**Table 2A**〉 Industry of Employment by Residential Location (percentage) 1980 (Top Entry) and 1990 (Bottom Entry)

	Location			
Industry	Central City	Old Manufact	Suburbs	Total
Manufact	31.2 17.6	53.1 27.4	34.3 27.3	38.8 25.5
Trade	19.9 19.9	13.3 21.5	17.5 19.1	16.8 19.9
Service	10.6 22.4	9.1 17.7	10.1 14.5	9.9 16.8
Others	38.2 40.2	24.6 33.3	38.2 39.2	34.4 37.8
Total	100	100	100	100

Others : This category includes construction, transportation, communications, public utilities, and public administration.

〈Table 2B〉 Earnings by Industry and Location 1980 (Top Entry) and 1990 (Bottom Entry) (Standard deviations are in parentheses)

	Location			
Industry	Central City	Old Manufact	Suburbs	Total
Manufact	32773 (12860) 28445 (12473)	35510 (9449) 29527 (9799)	35264 (11824) 29489 (12496)	34950 (11214) 29368 (11787)
Trade	27025 (13579) 21065 (8470)	30202 (11173) 24278 (12048)	33346 (15587) 28685 (19202)	31145 (14479) 26049 (16117)
Service	22656 (8542) 21665 (9608)	28104 (10746) 22971 (13177)	29817 (13783) 25074 (12678)	27829 (12377) 23662 (12200)
Others	31710 (10963) 27287 (11094)	33945 (12125) 26676 (9787)	36227 (12498) 34646 (20622)	34762 (12230) 31388 (17466)
Total	30145 (12355) 24997 (10983)	33748 (10794) 26285 (11195)	34746 (13139) 30716 (17719)	33539 (12497) 28516 (15350)

Annual wage and salary earnings (1989 dollars)
Others : This category includes construction, transportation, communications, public utilities, and public administration.

Our study differs from other recent research in that we categorize individuals by residential location, rather than job location. This is partly motivated by the assumption that an individual's residential location restricts his access to jobs. To verify this assumption, we examine the relationship of job location to residential location. The 1980 PUMS data contain job location information for sub-county level (CGRP) in Allegheny County while the 1990 PUMS do not.[11] <Table 3> gives job location information for 1980 data. For each cell of the table, the first row is for the number of individuals, the second row is the percentage of the total sample, the third row is the percentage in each job location of those in the specified residential location, and the fourth is the percentage in each residential location with a given job location. Approximately two-thirds of the central city and old manufacturing area residents work in their own residential areas, although less than half of the suburban residents work in the suburbs. Clearly a substantial number of workers commute to jobs outside of their own areas, although the fact that fewer workers commute out of declining areas than out of expanding areas suggest that there may be some barriers to central city and old manufacturing area residents.

V. Regression Specification and Results

The natural logarithm of hourly wage is used as a dependent variable in the earnings equation. The explanatory variables for the earnings equation are age, age square, years of schooling, and marriage (spouse present). The interpretation of marriage in the wage equation is the subject of some debate. The commonly found positive coefficient may represent returns to stability that are valued both in the labor market and in marriage. The explanatory variables for the location choice equation are the same variables in earnings equations as well as variables describing family composition which are assumed to affect location choice but not earnings.[12] Those variables are the number of household members, a dummy variable indicating the presence of children aged 6 to 17 years, and a dummy variable indicating the presence of children in both the 0 to 5 and 6 to 17 age ranges.

<Table 4> presents the results of the location choice probit model.[13] The dependent

〈Table 3〉 Number of Workers by Residential Location and Work Location in 1980

		Job Location				Total
		Central City	Old Manufact	Suburbs	Others	
Central City	N Percent Row percent Col percent	272 13.77 66.34 34.69	37 1.87 9.02 7.69	83 4.20 20.24 14.54	18 0.91 4.39 12.95	410 20.76
Old Manufact	N Percent Row percent Col percent	107 5.42 19.60 13.65	367 18.58 67.22 76.30	40 2.03 7.33 7.01	32 1.62 5.86 23.02	546 27.65
Suburbs	N Percent Row percent Col percent	405 20.51 39.74 51.66	77 3.90 7.56 16.01	448 22.68 43.96 78.46	89 4.51 8.73 64.03	1019 51.59
Total	N Percent	784 39.70	481 24.35	571 28.91	139 7.04	1975 100.00

〈**Table 4**〉 Location Choice Regression Results

	Location Choice			
	1980		1990	
	Subs - Central	Subs - Old Man	Subs - Central	Subs - Old Man
Age	0.037 (0.030)	0.025 (0.028)	0.104 (0.046)**	0.104 (0.042)**
Agesq	-0.394 (0.362)	-0.307 (0.338)	-1.100 (0.559)**	-1.161 (0.515)**
Schooling	0.110 (0.017)*	-0.014 (0.018)	0.117 (0.043)*	0.026 (0.041)
Marriage	0.316 (0.086)*	0.122 (0.081)	0.364 (0.110)*	0.134 (0.103)
Num. Household	-0.097 (0.022)*	-0.005 (0.021)	-0.159 (0.037)*	-0.058 (0.034)
Old Children	0.156 (0.066)**	-0.086 (0.062)	0.100 (0.098)	-0.039 (0.085)
Young Children	0.148 (0.093)	-0.036 (0.088)	0.254 (0.148)	-0.112 (0.131)
Constant	-1.598 (0.604)*	0.007 (0.566)	-2.834 (0.999)*	-1.959 (0.908)**
N	3192	3511	1588	1753

Dependent Variable: =1 if suburaban residence, =0 otherwise
Age : age in years
Agesq : age*age/1000
Schooling : years of education
Marriage : =1 if spouse are present, =0 otherwise
Num. Household: number of household members
Old Children : =1 if only has children who are 6 to 17 years old, =0 otherwise
Young Children : =1 if some children are 6 to 17 years old and some are under 6 years old, =0 otherwise
* : significant at 1 percent level
** : significant at 5 percent level
Standard errors are in parentheses.

schooling and marriage and the concave relationship with age are similar to patterns usually estimated in wage equations. This suggests that characteristics which lead to high wages also lead to the choice of a suburban residence, as would be indicated by the selective migration model. However, the significance of these coefficients is low by comparison with earnings equations. Likewise, the family structure variables are only occasionally significant. The point estimates indicate that large households prefer the city, and the presence of children leads to a preference for the old manufacturing area.

<Table 5A> and <Table 5B> provide regression results of earnings equations for 1980 data and for 1990 samples, respectively. The first row of each cell provides the results of endogenous location model and the second row contains results which treat location as

〈Table 5A〉 Earnings Regression Results in 1980 Top entry :sample selection model estimates Bottom entry : OLS estimates

	Location			
	Central City	Old Manufact	Suburbs	
Age	0.039 (0.016) 0.034 (0.014)	0.024 (0.011) 0.024 (0.011)	0.054 (0.008) 0.052 (0.008)	0.053 (0.008)
Agesq	-0.353 (0.190) -0.306 (0.177)	-0.233 (0.137) -0.228 (0.136)	-0.557 (0.100) -0.540 (0.100)	-0.541 (0.097)
Schooling	0.059 (0.012) 0.042 (0.009)	0.032 (0.007) 0.034 (0.008)	0.054 (0.006) 0.045 (0.006)	0.045 (0.005)
Marriage	0.209 (0.042) 0.182 (0.038)	0.117 (0.032) 0.110 (0.031)	0.087 (0.025) 0.074 (0.025)	0.076 (0.027)
Constant	0.487 (0.328) 0.555 (0.305)	1.243 (0.323) 1.113 (0.233)	0.170 (0.198) 0.390 (0.170)	0.366 (0.243)
σ	0.420 (0.039)	0.350 (0.047)	0.356 (0.013)	0.343 (0.020)
ρ	0.532 (0.181)	0.354 (0.519)	0.404 (0.157)	0.123 (0.863)
$\bar{\lambda}$	-1.180	-1.066	0.458	0.562
Adj R^2	0.09	0.04	0.08	
N	893	1212	2299	

Dependent Variable: log hourly wage
Age : age in years
Agesq : age*age/1000
Schooling : years of education
Marriage : =1 if spouse present, =0 otherwise
σ : estimates of standard deviations of the earnings equations
ρ : estimates of correlation between the errors of the location choice equation and earnings equation
$\bar{\lambda}$: average of λ (selection correction term)
Standard errors are in parentheses.

〈**Table 5B**〉 Earnings Regression Results in 1990 Top entry :Sample Selection model estimates Bottom entry : OLS estimates

	Location			
	Central City	Old Manufact	Suburbs	
Age	0.062 (0.029) 0.048 (0.026)	0.100 (0.025) 0.081 (0.021)	0.056 (0.017) 0.050 (0.016)	0.059 (0.017)
Agesq	-0.546 (0.354) -0.399 (0.316)	-1.051 (0.298) -0.862 (0.254)	-0.479 (0.205) -0.414 (0.194)	-0.510 (0.208)
Schooling	0.080 (0.031) 0.059 (0.024)	0.074 (0.021) 0.065 (0.022)	0.099 (0.016) 0.090 (0.016)	0.093 (0.016)
Marriage	0.177 (0.058) 0.159 (0.055)	0.079 (0.051) 0.079 (0.048)	0.115 (0.033) 0.103 (0.035)	0.105 (0.033)
Constant	-0.016 (0.611) 0.301 (0.566)	-0.373 (0.533) -0.255 (0.472)	-0.270 (0.414) 0.061 (0.361)	-0.285 (0.398)
σ	0.460 (0.061)	0.505 (0.052)	0.442 (0.016)	0.455 (0.018)
ρ	0.447 (0.353)	0.688 (0.127)	0.383 (0.189)	0.475 (0.140)
$\bar{\lambda}$	-1.245	-1.112	0.409	0.519
Adj R^2	0.12	0.09	0.11	
N	393	558	1195	

Dependent Variable: log hourly wage
Age : age in years
Agesq : age*age/1000
Schooling : years of education
Marriage : =1 if spouse present, =0 otherwise
σ : estimates of standard deviations of the earnings equations
ρ : estimates of correlation between the errors of the location choice equation and earnings equation
$\bar{\lambda}$: average of λ (selection correction term)
Standard errors are in parentheses.

〈**Table 6A**〉 Decomposition of Wage Differences in 1980

Location Pair	Place Effects	Sorting Effects		Total
	$(\beta_s-\beta_{cm})\bar{X}_{cm}+(\sigma_{su}-\sigma_{cmu})\bar{\lambda}_{cm}$	$\beta_s(\bar{X}_s-\bar{X}_{cm})$	$\sigma_{su}(\bar{\lambda}_s-\bar{\lambda}_{cm})$	$E(W_s \mid \bar{X}_s, U>0)-E(W_c \mid \bar{X}_{cm}, U\le 0)$
Subs - Central City	-0.133 0.106	0.027 0.024	0.236 ·	0.130 0.130
Subs - Old Manufact	-0.060[a] 0.009	0.001 0.000	0.069 ·	0.010 0.010

〈Table 6B〉 Decomposition of Wage Differences in 1990

Location Pair	Place Effects	Sorting Effects		Total
	$(\beta_s-\beta_{cm})\bar{X}_{cm}+(\sigma_{su}-\sigma_{cmu})\bar{\lambda}_{cm}$	$\beta_s(\bar{X}_s-\bar{X}_{cm})$	$\sigma_{su}(\bar{\lambda}_s-\bar{\lambda}_{cm})$	$E(W_s\mid\bar{X}_s, U>0)-E(W_c\mid\bar{X}_{cm}, U\leq 0)$
Subs - Central City	-0.158[a] 0.126	0.048 0.044	0.280 ·	0.170 0.170
Subs - Old Manufact.	-0.249 0.106	0.026 0.024	0.352 ·	0.130 0.130

Tob entry : sample selection estimates
Bottom entry : OLS estimates
$\beta_s(\bar{X}_s-\bar{X}_{cm})$: wage differences due to observed characteristics differences
$\sigma_{su}(\bar{\lambda}_s-\bar{\lambda}_{cm})$: wage differences due to unobserved characteristics differences
a : indicates selection effects are not significant.

〈Table 7〉 Earnings and Time to work by Residential Location and Work Location in 1980

		Job Location				
		Central City	Old Manufact	Suburbs	Others	Total
Central City	N	272	37	83	18	410
	Wage($)	28834	35493	28543	34287	29616
	Time(min.)	21	30	26	58	24
Old Manufact	N	107	367	40	32	546
	Wage($)	31226	34421	34694	31984	33672
	Time(min.)	30	16	28	47	22
Suburbs	N	405	77	448	89	1019
	Wage($)	36138	37973	33191	38931	35225
	Time(min.)	30	29	19	43	26
Total	N	784	481	571	139	1975
	Wage($)	32933	35072	32612	36730	33631
	Time(min.)	27	19	21	45	24

Others : out of Alleghenev County
Wage : yearly wage (real term in 1989 $)
Time : time to work (minutes)

exogenous. In the endogenous location regression results for the suburbs, the first column provides estimates corrected for the suburbs-central city choice and the second column provides estimates corrected for the suburbs-old manufacturing area choice.
variable is 1 for suburban residence and 0 otherwise. The positive coefficients on

Turning first to the coefficients on the sample selection terms, the estimates decompose the covariance of the location choice and earnings errors into the variance of the earnings error and the correlation of the errors.14) In all four cases, the correlation between the location choice error and the earnings error is positive. This is consistent with the selective migration model but not consistent with the Roy model of comparative advantage. These correlations are significantly different from zero in all cases except the equations which control for the choice between residence in the suburbs and the old manufacturing area in 1980. This, combined with the poor fit of the probit for this choice, suggests that these two locations were very similar in 1980.

Another piece of evidence that supports the selective migration model is the attenuation of the other coefficients when location is assumed to be exogenous. In the cases where the sample selection coefficients are significant, the coefficients on age, age squared, education and marriage have larger magnitudes in the sample selection models than in their OLS counterparts. This suggests that location choice is serving to truncate the samples in a manner that is correlated with income.

Table 6 decomposes the spatial variation in average earnings into place effects and sorting on observable and unobservable characteristics. The bottom entry in each cell is for the model in which location choice is assumed to be exogenous. In each of the location pairs, most of the difference in average earnings is attributed to a more generous earnings equation in the suburbs. Differences in individual characteristics play a relatively minor role.

A much different picture is presented by the top entry of each row which are calculated from the estimates which allow for location choice. In each case, the estimates suggest individuals who live in the suburbs have, on average, much higher levels of unobserved productive attributes. In fact, in each case, these differences would lead to earnings differences higher than those actually observed. The estimates indicate that the positive difference in attributes is partially offset by negative place effects -- identical individuals would actually have lower earnings in the suburbs. For example, the apparent 13% higher average earnings in the suburbs when compared to the old manufacturing area in 1990, decomposes into a 25% earnings disadvantage in the suburbs for someone who shares the attributes of the average old manufacturing area resident and a 38% earnings advantage due to more productive attributes of the suburban resident. The exception to this pattern is the comparison of the suburbs and the old manufacturing area in 1980 at which time the average wage difference was negligible.

These results are clearly contrary to a spatial mismatch model. Rather than attributing high suburban wages to a premium for otherwise identical workers, the results suggest that suburban workers are much more productive and that they, in fact, are paid less than equivalent workers in the old manufacturing (1990) or city (1980 and 1990).

How can the estimated wage premium for city and old manufacturing area residents be explained? One possible explanation is that it represents a compensation for disamenities of these areas. People are willing to live in the suburbs, even though they earn less, because of the positive utility from public goods available in the suburbs.

VI. Conclusion

The long trend of deindustrialization and job suburbanization in the U.S. economy has continued during 1990s. The spatial variation in economic growth due to the economic restructuring process affects spatial earnings distributions. Although many spatial mismatch studies exist for blacks, spatial effects for whites have been much less studied. In this paper, the spatial earnings variation of low skilled white males are examined. By allowing more than two locations and the endogeneity of location choice, we overcome some of the limits of traditional spatial mismatch research.

Descriptive statistics indicate that the suburbs-central city earnings gap increased and the earnings of residents in old manufacturing communities declined dramatically during 1980s. Estimates which do not account for the endogeneity of residential location suggest that the earnings gaps are due to wage premia in the suburbs. However, after accounting for the endogeneity of residential location, the earnings gaps between suburbs-central city and between suburbs-old manufacturing area are primarily due to differences in attributes of the residents.

The question remains whether more productive workers live in the suburbs in order to consume a more attractive bundle of amenities or in order to obtain access to expanding job sub-centers. Consider the indirect evidence regarding wage and travel time to work by residential and work locations presented in Table 7. In each cell, the first row is for the number of individuals, the second row is for annual wage (1989 dollars), and the third row is for travel time to work. This table shows that the patterns of work location choices are different among the residents of each location. Distinct from the residents of other areas, suburban residents who work where they live have considerably lower average wages than suburban residents who commute to either the city or the old manufacturing area. This is consistent with a model in which location choice is driven by locational differences in the marginal utility of earnings rather than by access to jobs.

Our estimates lead to the following conclusions. First, after accounting for endogenous
location, the wage premia are negative for the suburban area. This negative place effect may be due to the compensation for suburban amenities. Second, between 1980 and 1990, the attributes of the residents of the old manufacturing area declined substantially. This implies that the outmigration of individuals with productive attributes occurred during 1980s.

Although this work has focused on poorly educated white men, it has some implications for the understanding of the spatial mismatch hypothesis as it is commonly applied to African-Americans. Our results suggest removing suburban housing market barriers faced by blacks would not lead to an increase in earnings. Barriers to mobility lower the utility of blacks by preventing the consumption of amenities which are preferred by productive workers, but they do not appear to lower earnings.

Notes

1) For the review of spatial mismatch hypothesis, see Holzer(1991) and Kain(1992).
2) There are some spatial mismatch studies which refer to whites. However, those studies measure the spatial mismatch of blacks using comparisons between blacks and whites. For example, see McMillen(1993).
3) Calculated from County Business Patterns 1980 and 1990. For steel industry employment, we calculated the employment in 'Blast furnace and basic steel products' category (3 digit SIC code is 331).
4) According to American Statistics Abstract 1993, this ratio is the third highest among 20 largest metropolitan areas.
5) McMillen(1993), Engberg and Kim(1993), and Kim(1993) recognize the endogenous location choice.
6) The model is developed for two location choice case. Since we divide Allegheny County into three areas(central city, old manufacturing area, and expanding suburbs), the model is applied separately to pairs of areas. The potential problems of this are discussed in section V.
7) Our model is different from the Roy model(Roy, 1951) in which choice is based only on earnings potential. For example, in McMillen(1993), (work) location choice is decided by the potential earnings (minum commuting cost) in each location. However, in our model, the (residential) location choice depends on not only potential earnings in each location (W^*) but also the value of earnings (γ) and the amenities (A). The value of earnings and amenities may not be important when choosing work location, but it would be significant when choosing residential location. The location specific marginal utility of earnings and amenities lead to different implications between our residential location choice model and work location choice model based on the Roy model. For example, in our model, even though the

potential earnings after deducting commuting cost are higher in the central city, if the values of earnings (γ) and/or amenities (A) are higher in the suburbs, one would choose to reside in the suburbs rather than in the central city.

8) We deleted the individuals who reported zero years of education.
9) Hereafter, whenever we mention suburbs, we are referring to the expanding suburban areas which excludes the old manufacturing area. Because of differences in the sub-county area classifications between 1980 and 1990 data, the three areas defined in this paper are not exactly matched between 1980 and 1990 data. Some areas classified as central city in 1980 data are classified as suburbs in 1990 data (the population in these areas were 0.3% of Allegheny county population in 1980) and some areas classified as suburbs in 1980 data are classified as old manufacturing areas in 1990 data (the population in these areas were 1.7% of Allegheny county population in 1980).
10) In PUMS, the wage earned in 1979 is reported for 1980 data and the wage earned in 1989 is reported for 1990 data.
11) The 1980 PUMS data provide the job location information for a random subset which is about half of the total sample.
12) An exclusion restriction is not required in the wage equation; the coefficient on the sample selection term is identified from the non-linearity of the selection probit. However, exclusion restrictions can add considerably to the precision of the estimates.
13) The location choice equation and wage equation are estimated jointly for each location using an efficient maximum likelihood estimator rather than the more common but inefficient two step estimator. We also estimated the model using a linear probability model (Robinson, 1989) and using a multinomial logit trichotomous choice model (Lee, 1983). Although these variations place different restrictions on the joint distribution of the error terms, the results were qualitatively similar.
14) Note that $\sigma u=1$, thus $\sigma su=\sigma s^* \rho su$ and $\sigma cu=\sigma c^* \rho cu$.

References

Acs, Gregory and Sheldon Danziger (1993), "Educational Attainment, Industrial Structure, and Earnings: 1973-1987," *Journal of Human Resources* 28(3), 618-648.

Bound, John and Richard Freeman (1992), "What wend wrong? The Erosion of Relative Earnings and Employment Among Young Black Men in the 1980's," *Quarterly Journal of Economics* 107(1), 201-232.

Card, David and Alan Krueger (1992), "School Quality and Black-White Relative Earnings: A Direct Assessment," *Quarterly Journal of Economics* 107(1), 151-200.

Engberg, J. (1993), "Person or Place? Spatial Variation within Local Labor Markets," *Heinz School Working Paper* 93-35, Carnegie Mellon University.

Engberg, J. and T. Kim (1994), "Person of Place? Sources of Intra-Metropolitan Earnings Variation," presented at 1994 Winter Meeting of The American Real Estate and Urban Economics Association, Boston.

Heckman, J. (1979), "Sample Selection Bias as a Specification Error," *Econometrica* 47, 152-161.

Holzer, Harry J. (1991), "The Spatial Mismatch Hypothesis: What Has the Evidence Shown?," *Urban Studies* 28(1), 105-122.

Ihlanfeldt, Keith R. (1988), "Intra-Metropolitan Variation on Earnings and Labor Market Discrimination: An Econometric Analysis of the Atlanta Labor Market," *Southern Economics* Journal, July, 123-140.

Ihlanfeldt, Keith R. and David Sjoquist (1989), "The Impact of Job Decentralization on the Economic Welfare of Central City Blacks," *Journal of Urban Economics* 26, 110-130.

Ihlanfeldt, Keith R. and David Sjoquist (1990), "The Effect of Residential Location on the Probability of Black and White Teenagers Having a Job," *Review of Regional Studies*, 10-20.

Kain, John F. (1968), "Housing Segregation, Negro employment, and Metropolitan Decentralization," *Quarterly Journal of Economics*, May 82(2), 176-197.

Kain, John F. (1992), "The Spatial Mismatch Hypothesis: Three Decades Later," *Housing Policy Debate* 3(2), 371-460.

Katz, Lawrence and Kevin Murphy (1992), "Changes in relative Wages, 1963-87: Supply and Demand Factors," *Quarterly Journal* of Economics 107(1), 35-78.

Kim, J. (1993), "Divided Metropolis: Intra-metropolitan Spatial Segmentation of the Labor Market," Ph.D. Dissertation, The Heinz School of Public Policy and Management, Carnegie Mellon University.

Lee, L-F. (1983), "Generalized Econometric Models with Selectivity," *Econometrica 51*, 507-512.

Maddala, G.S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics* (New York, NY: Cambridge University Press).

McMillen, Daniel P. (1993), "Can Black Earn More in the Suburbs? Racial Differences in Intra-Metropolitan Earnings Variation," *Journal of Urban Economics* 33, 135-150.

Robinson, Chris (1989), "The Joint Determination of Union Status and Union Wage Effects: Some Tests of Alternative Models," *Journal of Political Economy* 97(3), 639-667.

Roy, A. D. (1951), "Some Thoughts on the Distribution of Earnings," *Oxford Economic Paper* 3, 135-146.