Institute for Research on Poverty
Discussion Paper no. 1089-96

PERSON OR PLACE?
Parametric and Semiparametric Estimates of Intrametropolitan Earnings Variation

John B. Engberg
and
Taeil Kim

The Heinz School of Public Policy and Management
Carnegie Mellon University

May 1996

We have benefitted from comments by Glen Cain, Guido Imbens, and from participants at the Carnegie Mellon/University of Pittsburgh Applied Microeconomics Workshop and the Institute for Research on Poverty Seminar. This research was supported by the U.S. Department of Health and Human Services through the Institute for Research on Poverty (Engberg) and a HUD Doctoral Research Grant (Kim). Points of view or opinions expressed in this paper are those of the authors and do not necessarily represent either the U.S. Department of Housing and Urban Development, the U.S. Department of Health and Human Services, or the Institute for Research on Poverty.

Abstract

Some scholars have attributed earnings differences among locations to labor market conditions ("place effects") whereas others have focused on the skill level of residents ("person effects"). We estimate a variety of selection models in an effort to detect differences in labor market conditions while controlling for differences in skill levels. We maintain the assumption that there are no barriers to mobility within a metropolitan area for highly educated white men, which implies that intra-urban differences for this group reflect sorting by skill and earnings rather than real wage differences for equally productive workers. This prediction allows us to reject several conventional parametric selection models. We estimate a semiparametric selection model that yields strong evidence that, for less educated white men, the apparent suburban earnings premium is due to sorting rather than labor market differences.

PERSON OR PLACE?
Parametric and Semiparametric Estimates of
Intrametropolitan Earnings Variation

INTRODUCTION

Is the relative demand for low-skill employment higher in the suburbs than in the city? If so, and if barriers to residential mobility and commuting restrict the suburban supply of low-skill workers, then low-skill suburban residents will earn more than comparable center-city residents.

Identifying comparably skilled workers is the greatest challenge in empirically verifying this hypothesis. Many dimensions of skill are not observable, making possible the alternative conclusion that observed earnings differences represent differences in “person” rather than in “place.” Urban economists have long recognized that residential location is endogenous. Housing filtering, zoning restrictions, and the positive income elasticity of lot size imply that people with high earnings are more likely to choose a suburban residence. Therefore, a suburban earnings premium for workers of apparently equal skill may in fact reflect compensation for unobserved skills and the resulting preference for a suburban residence. Consistent estimation of the impact of location on earnings requires accounting for the endogeneity of residential location.

The spatial labor market disequilibrium that some scholars posit for low-skill workers is not likely to be present for high-skill workers. The costs of commuting or relocating may create an effective barrier for many low-skill workers but are not likely to be influential in the labor market for high-skilled workers. We assume that any substantial earnings difference between high skilled suburban and city residents is due to residential choice rather than spatially segmented labor markets.

We frame the estimation of spatial earnings differences in the presence of residential endogeneity as a sample selection problem and examine a variety of specifications. We estimate the suburban earnings premium using a sample of employed prime-age white men from Allegheny County (Pittsburgh). We find that conventional parametric corrections for sample selection bias are not adequate to account for the suburban earnings premium of high-skilled workers. Therefore, we turn to a semiparametric sample selection estimation method. When combined with the assumption that residential choice accounts for the entire high-skill suburban premium, we reject the hypothesis of a suburban premium for low-skilled residents.

Social science research provides many theoretical and empirical examinations of intrametropolitan spatial earnings differences. The monocentric urban model implies that wages will decrease as distance between the place of employment and the central business district (CBD) increases (e.g., Mills and Hamilton, 1994). However, decentralization of production over the past several decades may weaken or reverse this prediction. Ihlanfeldt (1992) does not find negative wage gradients for several low-skill occupations. This finding corresponds to anecdotal and ethnographic evidence of suburban low-skill labor shortages.

Although the wage gradient is expressed in terms of place of employment, the spatial mismatch literature makes the important connection to place of residence.[1] Recent empirical work suggests that wage rates and employment rates are lower for youth who live far from areas of employment concentration or employment growth (Ihlanfeldt and Sjoquist, 1990; Raphael, 1995; O'Regan and Quigley, 1995), thus providing additional support for the existence of spatial disequilibrium.[2]

Although these empirical explorations are executed with great care, they may not fully account for endogenous location. Wage-gradient estimates based on place of work will be biased if there is a systematic relationship between distance from the CBD and skill requirements. For example, firms

[1]The spatial mismatch hypothesis (Kain, 1968) attributes racial differences in labor market outcomes to housing market barriers. Black workers are constrained to live in the city, preventing access to a spatially decentralized pool of low-skill jobs. See Holzer, 1991; Kain, 1992; Ihlanfeldt, 1994; and Jencks and Mayer, 1990, for recent reviews of the empirical literature.

[2]Alternatively, some have hypothesized that the demographic composition of the residential neighborhood has an important impact on labor market outcomes, especially for youth. See O'Regan and Quigley, 1995; Case and Katz, 1991; and Borjas, 1992, for empirical examinations and Manski, 1993, for a critique of this hypothesis.

farther from the CBD may be newer and employ technology that requires expertise that is not measured by standard surveys. The spatial-mismatch hypothesis estimates have focused on youth to minimize the problem of endogenous location, but the possibility of unmeasured family characteristics that are correlated to location and labor market outcome is a recurring concern.[3]

Although the preponderance of the empirical work in spatial earnings differences has focused on minority youth, we prefer to work with a sample of prime-age white men. We choose to focus on earnings, and implicitly on wages, leaving the analysis of spatial patterns of labor force participation and employment for future study. Excluding women and youth from the sample reduces the sample selection bias from the labor force participation and employment decisions. Studying spatial earnings patterns of white workers eliminates the possibility that spatial patterns of racial discrimination in the labor market are affecting the results.

Although the existence of spatial earnings differences among prime-age white men is of interest in its own right, the results provide a baseline with which to compare spatial patterns for other demographic groups. Furthermore, if workers from all demographic groups compete for the same jobs, then spatial wage differences cannot persist for some groups but not others. Even if (spatially uniform) labor market discrimination leads to racial wage premia, labor market competition will equate spatial wage premia for all demographic groups.[4]

We examine the Pittsburgh area because its economic and physical characteristics are conducive to spatial earnings differences. Like many metropolitan areas, economic growth in the

---

[3]For example, Ihlanfeldt (1988) rejects endogenous location using a parametric selection model, but McMillen (1993) finds strong evidence of sample selection in a model which assumes residential choice is entirely driven by labor market differences.

[4]Strazheim, in his frequently cited theoretical work (1980) examining racial differences in wage gradients, provides a model in which white workers have a negative wage gradient, but black workers have a positive wage gradient. However, this model rests on the premise of a very low elasticity of substitution between black and white workers. Such a premise does not have any theoretical or empirical support. See Engberg (1996) for a theoretical examination that permits greater substitution.

suburbs has greatly exceeded growth in the city, especially in industries and occupations with relatively low skill requirements. Furthermore, residential migration from the city to the suburbs has been hampered by a considerable migration out of the region, which has depressed housing prices in the central city. Commuting is notoriously difficult due to three major rivers and the ubiquitous hills. These features provide ideal conditions for a spatially segmented labor market. The observed 21 percent earnings suburban premium in the expanding suburbs relative to the city is consistent with this conjecture (Table 1). Finally, unlike many metropolitan areas, prime-age white men represent a very large portion of the labor supply at all skill levels in both the central city and the suburbs (Table 1).

The analysis is complicated by differences in economic growth among the suburbs. We separate the older suburbs and former steel mill towns in the southeastern portion of the county from the expanding suburbs. This permits the comparison of earnings between the center city and expanding suburbs. However, when modeling the residential location choice, we include all three locations.

This introduction is followed by a brief review of our methodological strategy. We then present a sequence of models of spatial earnings differences and the corresponding estimates. We conclude with some comments on the identifying assumptions of the final estimates and on the implications of our findings.

IDENTIFYING PLACE EFFECTS

Place effects in an earnings equation can be identified by controlling for differences in observed and unobserved average skill among locations. Consider an experiment in which individuals are randomly assigned to a residential location. Any resulting differences in average earnings would

TABLE 1

Average Weekly Earnings of Employed White Prime-Age Men
(Allegheny County, Pennsylvania)

Location	*Average Weekly Earnings*	*Percentage of All Employed Prime-Age Men Who Are White*
City of Pittsburgh	326	83.4
Expanding Suburbs	401	98.1
Older Suburbs and Steel Towns	370	91.0

Source: Authors' tabulation of 1980 Public-Use Microsample of the U.S. Census.

Notes: "Average Weekly Earnings" is a geometric mean. Population: Male workers between the ages of 25 and 55 who worked at least 35 hours per week, at least 40 weeks in 1979, and earned at least $100 per week.

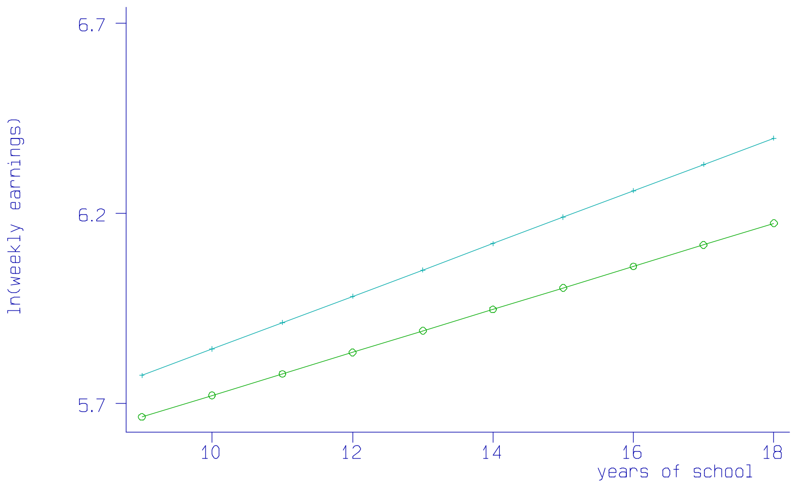
be due to the “treatment” associated with residential location. In the absence of a controlled experiment, place (treatment) effects can be identified only by modeling the process that assigns individuals to each location (Manski, 1993). Consistent estimates of wage differences due to place effects can then be obtained by accounting for the endogeneity of location in the wage formation process.

We formulate a general model in which residential location is determined by differences in potential wages and amenities in each location. We examine several parametric and semiparametric estimators based on decreasingly restrictive sets of assumptions. The goal is to find a model that confirms the maintained assumption that there are no place effects for highly educated workers.

Consider the panels of Figure 1. Panel A provides a stylized representation of the apparent earnings levels and returns to skill for the expanding suburbs and the central city. Separate ordinary least squares (OLS) regressions of earnings on education and other human capital measures demonstrate that earnings are higher in the suburbs and that the gap increases with education. (Regression coefficient estimates are in the appendix.) These estimates would be consistent if residential location is exogenous. However, the large difference between the city and expanding suburbs for highly educated workers leads us to reject the hypothesis of exogenous location. Endogenous location causes these equations to be biased, both in level and slope. If an estimation method that accounts for endogenous location is consistent, it should reveal lines that converge for highly educated workers. Therefore, any method that does not remove the spatial gap for highly educated workers is not properly specified and should not be used to estimate spatial differences for unskilled workers.

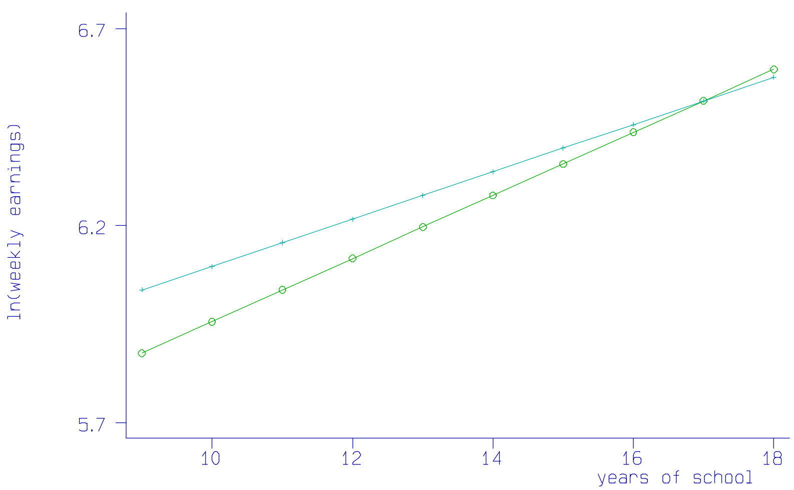
Unfortunately, none of the parametric estimators control for all the person effects among the highly educated workers. In order to obtain consistent estimates of the place effects for less educated workers, we use a method that provides consistent estimates of the location-specific earnings

FIGURE 1 A
CITY: --O-- SUBURBS: --+--



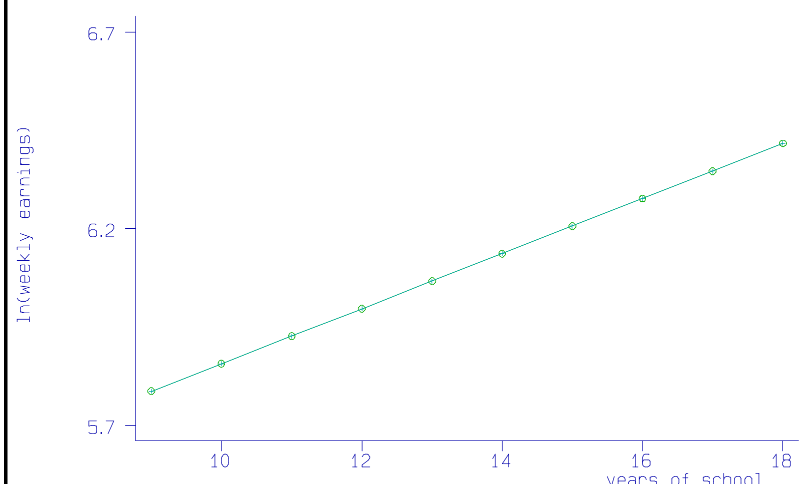
OLS ESTIMATES

FIGURE 1 B
CITY: --O-- SUBURBS: --+--



PLACE EFFECTS (HYPOTHETICAL)

FIGURE 1 C
CITY: --O-- SUBURBS: --+--



NO PLACE EFFECTS (HYPOTHETICAL)

regression slopes (but not the intercepts) under very unrestrictive assumptions (Ahn and Powell, 1993). We proceed by imposing the assumption of no place effects for highly educated workers, which allows us to calculate estimates of the place effects for less educated workers. If the resulting estimates look like Panel B in Figure 1, this will provide evidence of spatially segmented low-skill labor markets. But, if consistent estimates resemble Panel C, this will provide evidence that there are no place effects for low-skill workers.

This work draws on recent research in econometrics and in many applied fields which examines the impact of modeling assumptions on the estimation of treatment effects in self-selected populations. Mroz (1987), Newey, Powell, and Walker (1990), and Ahn and Powell (1993) examine progressively unrestrictive versions of Heckman's (1974) model of women's labor supply. Robinson (1989) examines the union wage effect. Heckman and Hotz (1989) and Heckman, Ichimura, Smith, and Todd (1995) analyze the impact of job training programs. Angrist (1995) examines the returns to schooling. Manski, Sandefur, McLanahan, and Powers (1992) analyze the impact of family structure on educational attainment. All of these applications require an understanding of a selection process in order to estimate the treatment effect of interest.

<u>Wage Differences by Education Level</u>

Row 1 of Table 2 presents the average earnings difference between the city of Pittsburgh and its expanding suburbs by education level. Clearly, the overall earnings difference of 21 percent implied by Table 1 is not accounted for by differing levels of education in the two locations. The differences increase with education level, with the exception of the most highly educated group.[5] Workers with 16 years of education (bachelor's degree) earn almost a third more if they live in the suburbs than if they live in the city. The greater differences for the more highly educated workers are

[5]Given the large sample size, these differences are estimated very precisely. For each category, the t-statistic is greater than 5.

TABLE 2

**Location Earnings Differences:
Suburban Premium by Education Category**

	Years of Schooling				
	8–11	12	13–15	16	16–20
(1) Raw Earnings Difference[a]	11.0	14.4	20.7	33.1	24.7
(2) Control for Exper, Edu, Marr[b]	11.3	9.3	9.4	18.5	17.4
(3) Exogenous Location[c]	6.7	11.7	13.3	15.1	17.7
(4) Para Error, Linear Choice[d]	136.0	142.0	144.0	149.0	155.0
(5) Para Error, Nonlinear Choice[e]	19.8	23.2	24.1	24.7	25.7
(6) Nonparametric Selection[f]	-5.4	-5.8	-5.6	-2.9	***
(bootstrap standard errors)	(8.3)	(4.9)	(3.2)	(1.8)	
Sample: City	317	695	262	227	266
Expanding Suburbs	476	1874	799	989	673

Source: Authors' tabulation of 1980 Public-Use Microsample of the U.S. Census.

Notes: Regression coefficients for rows 3–6 are presented in the appendix.

Rows 3–6: Table entries = $\overline{X}(\beta_s - \beta_c)$ where $\overline{X}$ is the mean by education category of the exogenous variables for the entire county; β_c and β_s are estimated coefficients using city and expanding suburbs workers, respectively.

[a]Percentage difference between residents of the expanding suburbs and the city of Pittsburgh.
[b]Table entries are coefficients on interactions of dummies for each education group with a dummy for the expanding suburbs in a regression that controls for experience, experience squared, years of education, and marital status.
[c]Estimated from OLS regression (equation (1)).
[d]Estimated from OLS regression that includes sample selection term defined in equations (4) and (5).
[e]Estimated from OLS regression that includes sample selection term defined in equations (6) and (7).
[f]Estimated from OLS regression on differences between individuals with similar predicted location choice probabilities. See equations (11) and (12).

Sample: Male workers between the ages of 25 and 55 with 8 years of schooling or more who worked at least 35 hours per week, at least 40 weeks in 1979, and earned at least $100 per week.

consistent with the hypothesized greater mobility of these workers and the positive relationship between income and preference for suburban location.

It is possible, however, that differences in observed (exogenous) characteristics other than education account for the earnings differences. Table 3 indicates that the average education, average age, and proportion married are higher in the expanding suburbs than in the city. These characteristics are usually associated with higher wages, and so might explain the observed earnings differences. These patterns are consistent with the theory of person effects: individuals with observed characteristics indicative of higher skills are more likely to live in the suburbs.

Row 2 of Table 2 reports the earnings differences after controlling for age, education, and marital status. As expected from the patterns of age and marital differences across locations, the estimated earnings differences are lower after controlling for these factors. However, there remains an unexplained earnings difference of over 15 percent for the highest education categories.[6] This suggests that sorting on observables is not sufficient to explain earnings differences between locations. Modeling of the location choice process is necessary to capture the dependence of location on unobservable characteristics that are correlated with earnings.

LOCATION CHOICE MODEL

The location choice model is a variation on the Roy model (Roy, 1951; Heckman and Honore, 1990), which has been used extensively to model earnings for self-selected populations. The Roy model is based on the idea that individuals choose the alternative that maximizes their earnings or utility. Therefore, individuals who are observed in a particular alternative are not a random sample of the underlying population, but will have *unobserved* characteristics that systematically compensate for

---

[6]Given the large sample size, these differences are estimated very precisely. For each category, the t-statistic is greater than 5.

TABLE 3

Average Characteristics by Location of Full-Time Prime-Age White Male Workers (Allegheny County, Pennsylvania)

Location	*Average Education*	*Average Age*	*Percentage Married*	*Sample Size*
City of Pittsburgh	13.3	37.7	.69	1790
Expanding Suburbs	13.7	39.2	.85	4849
Older Suburbs and Steel Towns	13.2	39.0	.81	2200

Source: Authors’ tabulation of 1980 Public-Use Microsample of the U.S. Census.

Notes: Population: Male workers between the ages of 25 and 55 who worked at least 35 hours per week, at least 40 weeks in 1979, and earned at least $100 per week.

levels of *observed* characteristics that predispose them against the chosen alternative. In the simple case with one observed characteristic, OLS regression using a selected sample will give a downward-biased estimate of the actual impact of the observed characteristic on the outcome. For example, the returns to education would be underestimated using either a sample composed only of high-income individuals or only of low-income individuals.

In the current application, a set of equations describes potential earnings in each of three locations.[7] An additional set of equations governs the choice of residential location. The two parts of the model are linked in that potential earnings is one of the factors that determines location choice.

Potential (or latent) earnings of person i in location j, W^*_{ij}, is given by the following equation:

$$W^*_{ij} = X_i\beta_j + \varepsilon_{ij} \tag{1}$$

Earnings depend on observable characteristics, X_i, according to the coefficients β_j, which may vary by location. For example, if less educated workers get paid more in the suburbs than in the city, this implies greater returns to education in the city than in the expanding suburbs. In addition to education, X_i includes experience, experience squared, and an indicator of marital status. The error term ε_{ij} represents unobserved characteristics that affect earnings for person i in location j. It is assumed to be independent of X_i in the population. As discussed above, if comparisons of potential earnings affect location choice, ε_{ij} will not be uncorrelated with X_i within a location.

The utility, U^*_{ij}, received from choosing a particular residential location is determined by the value of earnings that can be attained and by the value of amenities in the location:

$$U^*_{ij} = W^*_{ij}\gamma_j + X_i\alpha_j + v_{ij} \tag{2}$$

[7]See McMillen, 1993, for an earlier application of the Roy model to spatial earnings differences.

The coefficient γ_j is included to capture the idea of the Tiebout hypothesis (1956). The Tiebout hypothesis suggests that locations differ in the bundles of public goods provided to the residents. A location with a relatively large value of γ is more highly valued by high-income individuals. (The coefficient γ is assumed to be greater than or equal to zero in all locations.) Similarly, locations differ in their attractiveness to individuals of differing education, age, and marital status. These differences are captured by variation among locations in the parameter α.[8] The error term v_{ij} captures unobserved characteristics that have an impact on location-specific utility. As with the earnings error terms, v_{ij} is assumed to be independent of X_i in the population.

The individual is assumed to choose the location that provides the highest utility. This choice is indicated by a categorical variable, d, which equals the index of the chosen location. Earnings are observed for only the chosen location, W_{id}.

The endogeneity of location arises because, in general, there is a correlation between the error term of the earnings equation and the explanatory variables in the earnings equation for individuals who choose a location; that is, $E(\varepsilon_{ij}|X_i,d=j)$ varies with X_i. For example, if a location is valued by individuals with high earnings, then less educated individuals will chose that location only if they have unusually high unobserved characteristics that contribute to earnings in that location.

The concepts of place and person effects can be expressed as variation across location in particular parameters of the model. As mentioned above, place effects are captured by differences among locations in the earnings equation coefficients β_j. Person effects can arise from several sources. Spatial variation in the utility value of earnings (γ_j) indicates that some locations will attract high-earnings individuals and other locations will attract low-earnings individuals. Spatial variation in the wage error term (ε_{ij}) for each person implies that individuals choosing each location have particular

[8]We use the standard specification of experience (i.e., age-education-6) and its square in the wage equation, but use age and its square in the location-choice equation. The presentation ignores this slight difference in the explanatory variables.

skills that are demanded in that location. Correlation between the earnings and location error terms ($E(\varepsilon_{ij}v_{ij}) \neq 0$) will also lead to correlation between the error of the wage equation and the explanatory variables in the chosen location, although in this case earnings do not "cause" location.

Exogenous Location Choice

Substituting the location-specific earnings equation (1) into the location-specific utility equation (2) provides a reduced-form expression of the location specific utility:

$$U^*_{ij} = X_i(\beta_j\gamma_j + \alpha_j) + \varepsilon_{ij}\gamma_j + v_{ij} \qquad (3)$$

This representation suggests two special cases in which there are no person effects, i.e., cases in which location is exogenous with respect to earnings. If earnings do not have an impact on location choice (γ_j=0 for all j) then the composite error term ($\varepsilon_{ij}\gamma_j + v_{ij}$) from equation (3) is independent of the error term ε_{ij} in the wage equation. Similarly, if neither the utility of earnings nor the earnings equation error term vary by location (i.e., $\gamma_j=\gamma$ and $\varepsilon_{ij}=\varepsilon_i$ for all j), then the earnings equation error will not play a role in determining location. Note that in both cases the errors in the location choice equation (v_{ij}) must also be assumed independent of the earnings equation error terms (ε_{ij}).

Although these are very restrictive models, they are necessary to justify OLS estimation of location-specific earnings equations. Row 3 of Table 2 presents the place effects estimated under the assumption of exogenous location choice. (The notes to Table 2 provide the formula for place effects. The appendix contains the estimated regression coefficients.) The estimated place effects are increasing with education, ranging from 6.7 percent for the least educated to almost 18 percent for the most educated. The large estimated place effects for highly educated white men provide evidence against the hypothesis of exogenous location.[9]

[9]Of course, it is possible that the model is correct and the estimated place effect for the high-education groups is a consistent, but very imprecise, estimate of zero. This would not alter the

<u>Linear Location Choice with Parametric Error Distributions</u>

A conventional approach to controlling for selection effects in a model with multiple alternatives is to estimate the choice equation with multinomial logit (Lee, 1983). The MNL estimator is based on the assumption that the composite error term of the choice equation ($\varepsilon_{ij}\gamma_j+v_{ij}$) is independently and identically distributed across individuals and locations, with a type I extreme-value distribution. This permits the conditional expectation of the location-specific earnings equation error term to be written as a function of the predicted probability of living in the chosen location:

$$E(\varepsilon_{ij}|X_i,d=j) \;=\; \gamma_j\ \sigma_\varepsilon^2\ \lambda_j(X_i) \;=\; \gamma_j\ \sigma_\varepsilon^2\ \frac{\phi(\Phi^{-1}[P_j(X_i)])}{P_j(X_i)} \tag{4}$$

where:

$$P_j(X_i) \;=\; \frac{e^{X_i[\beta_j\gamma_j\,+\,\alpha_j]}}{\sum_{k=1}^{J} e^{X_i[\beta_k\gamma_k\,+\,\alpha_k]}} \tag{5}$$

The standard normal probability distribution function and cumulative distribution function are represented by ϕ and Φ, respectively. The earnings equation error term (ε_{ij}) is assumed to be normally distributed with variance σ_ε^2 and to be independent of v_{ij}.

There are two drawbacks to this specification. First, it assumes a functional form for the explanatory variables in the location-choice equation. Although equation (2) assumes that location choice depends on the explanatory variables through a linear index ($X_i\alpha_j$), there is little theory to justify this specification. However, we temporarily assume that the choice equation is linear in education, age, age squared, and marital status. Second, even if linearity in the explanatory variables is correct, the

---

conclusion that this model is not useful for estimating place effects. Given the difficulty of calculating either analytic or bootstrap standard errors for these estimates, it does not seem worthwhile to investigate further whether it is a question of bias or precision.

transformations in equations (4) and (5) from the linear index ($X_i\alpha_j$) to obtain the conditional expectation of the location-specific earnings error term rely on specific distributional forms for the wage and location-choice error terms ε_{ij} and v_j. Again, there is no theoretical justification for these forms. The estimated place effects of more than 100 percent reported in Row 4 of Table 2 suggest that these restrictions, although conventional, are not appropriate.

Nonlinear Location Choice with Parametric Error Distributions

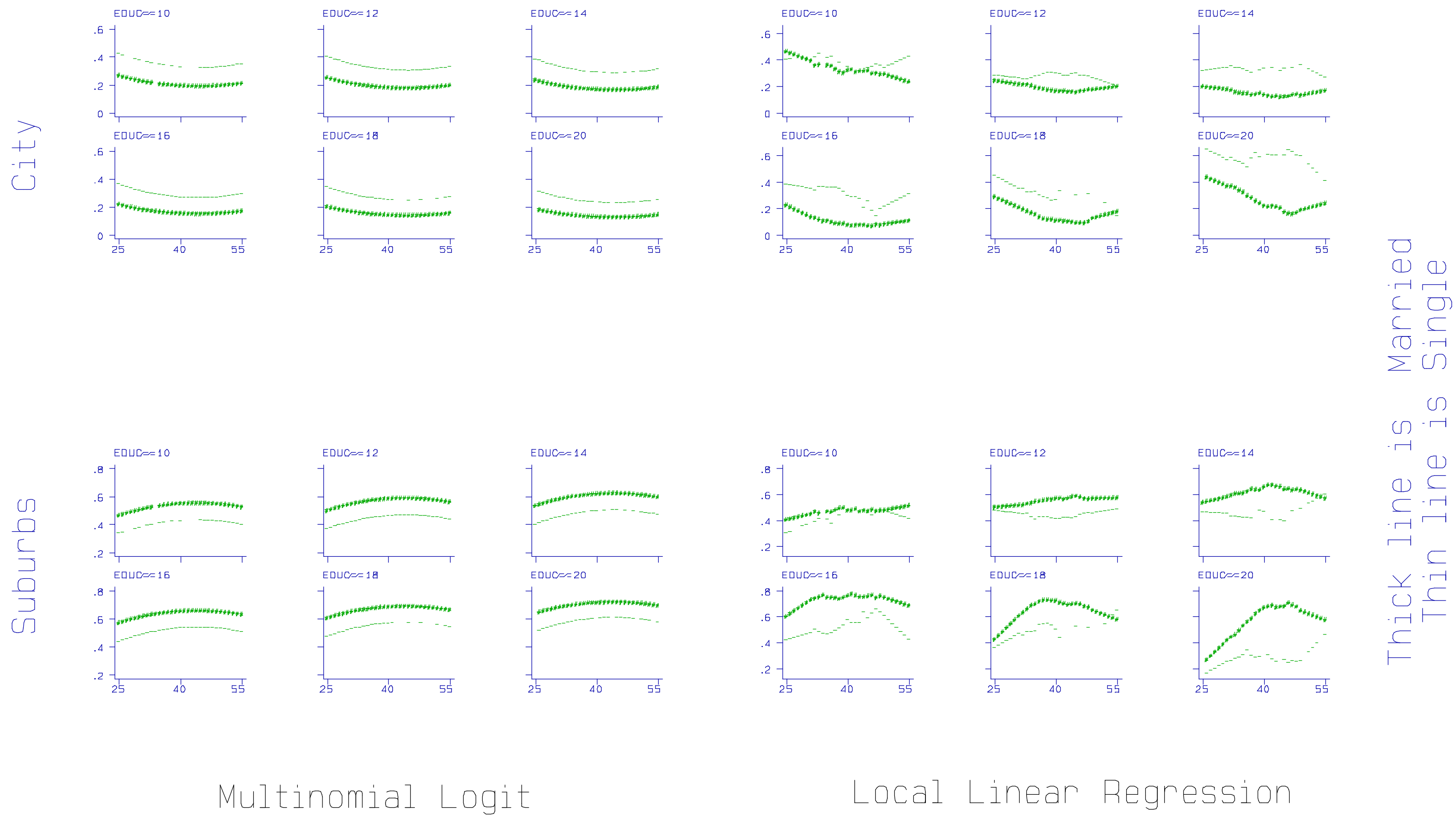
The assumption that the explanatory variables enter the location choice index with a particular functional form (i.e., linearity) can be relaxed by estimating the location choice equation nonparametrically. By changing $X_i\alpha_j$ to an unknown function $g_j(X_i)$, we allow for arbitrary nonlinearities, including interactions, of the explanatory variables. We estimate $P_j(X_i)$, j=1,2, using local linear regression.[10] Figure 2 presents the predicted probabilities from MNL (left side) and local linear (right side) regression. For example, the local linear estimates capture the nonlinearity that arises from the tendency of single middle-aged men to live in the city.

Not only do the nonlinearities captured by the local linear regression portray the choice process more accurately, they also help distinguish the impact of person from that of place. In the i.i.d. parametric model above, the person effects were captured by selection correction term $\lambda_j(X_i)$. Variation in this term (independent of X_i) depends entirely on the assumed distribution of the error

[10]We split the sample into ten cells defined by marital status and education category. For each observation, we estimate an OLS regression of location choice on age and years of education using all observations from the same cell whose age differs by no more than six years. (For observations whose age is less than six years from the minimum or maximum age, we use a regression centered seven years from the bound.) The predicted probability for each observation is calculated from their own regression. The bandwidth of +/- six years was chosen by leave-one-out cross validation (Hardle, 1990).

Alternative nonparametric methods involve adding polynomials in X to the choice equation (Mroz, 1987) or kernal estimation (Ahn and Powell, 1993). As pointed out by Heckman et al. (1995), local linear regression has the advantage of retaining prediction precision in sparsely populated portions of the distribution of explanatory variables. (Also see Fan, 1993.) See Heckman et al. (1995) for an application of local linear regression in the second step of a two-step selection model.

FIGURE 2



Predicted Location Probabilty v. AGE

terms. The nonlinear functions of X implicit in the local linear regression act as instrumental variables that provide additional independent variation in $\lambda_j(X_i)$.

Another possible source of independent variation in the selection correction term would be to include variables in equation (2) which affect location choice but are not correlated to the error term in the earnings equation. Unfortunately, variables that are likely to affect location choice, such as number of children, spouse's characteristics, and nonlabor income, are likely to be related to earnings. Therefore, we prefer to use nonlinearities for identification.

Replacing the linear MNL probability estimate in equation (4) with the local linear choice probability estimates retains the functional form of the error distributions while relaxing the linearity assumption on the explanatory variables:

$$E(\varepsilon_{ij}|X_i, d=j) = \gamma_j \; \sigma_\varepsilon^2 \; \lambda_j(X_i) = \gamma_j \; \sigma_\varepsilon^2 \; \frac{\phi(\Phi^{-1}[P_j(X_i)])}{P_j(X_i)} \tag{6}$$

where:

$$P_j(X_i) = \frac{e^{X_i\beta_j\gamma_j + g_j(X_i)}}{\sum_{k=1}^{J} e^{X_i\beta_k\gamma_k + g_k(X_i)}} \tag{7}$$

Row 5 of Table 2 presents the estimated place effects from this specification. Although they are considerably more reasonable than the previous row, the estimate of a 26 percent wage difference for the most educated group casts doubt on the assumed error distributions.

<u>Nonparametric Selection</u>

Unfortunately, removing the parametric distributional assumptions of the error terms does not yield tractable estimators of the form described above. The assumptions about the shape of the error distributions allow the expected value of the earnings error term to be expressed as a single known function of the estimated location choice probability. However, we can remove these restrictive assumptions at the cost of using an estimator that expresses the expectation of the earnings error term as an unknown function of the location choice probability:

$$E(\varepsilon_{ij}|X_i, d = j) = K[P_j(X_i)] \tag{8}$$

A sufficient condition for substituting a function of the scaler $P_j(X_i)$ for a function of the high dimensional vector X_i is that the reduced-form choice equation (i.e., choosing the maximum U^*_{ij}, as defined by equation (3), over all 3 locations) can be expressed as a dichotomous choice between the chosen location and all other locations. If the dichotomous choice equation can be represented by a function of X_i and an additive error term that is independent of X_i, then no further restrictions are necessary regarding the functional form for the explanatory variables or the distribution of the error term.[11]

We do not attempt to estimate the unknown function $K(P_j)$, but instead treat it as a nuisance parameter. Following a method developed by Ahn and Powell (1993), the slope coefficients of the earnings equation can be estimated by OLS after differencing to eliminate $K(P_j)$. The differences are taken pair-wise between individual a and individual b in location j who share similar values of the location-choice probability but differ in their values of X:[12]

$$W_{aj} - W_{bj} = (X_a - X_b)\beta_j + K[P_j(X_a)] - K[P_j(X_b)] + \eta_{aj} - \eta_{bj} \tag{9}$$

[11]This is known as index sufficiency and, in the statistical literature, is closely related to the use of propensity scores. Heckman et al., 1995, contains an illuminating presentation of index sufficiency. Use of propensity scores is due to Rosenbaum and Rubin, 1983. Angrist, 1995, ties these literatures together and presents alternative conditions that permit this substitution.

[12]Careful examination of the right side of Figure 1 indicates which pairs of values of X have the same predicted probability. For estimation purposes, all possible N*(N-1) pairs were weighted by a normal kernal evaluated at their probability difference. See Ahn and Powell (1993). Bandwidth for the kernal weight is selected by leave-one-out cross validation. Only pairs that are “close” in probability are left out and the chosen bandwidth minimizes the unweighted mean squared error of this subsample.

such that

$$X_a \neq X_b \ and\ P_j(X_a) = P_j(X_b) \tag{10}$$

The new error term (η_{aj}-η_{bj}) is uncorrelated with X_{aj}-X_{bj} by construction. This provides consistent estimates of the returns to education, experience, and marital status in each of the locations, without placing any restrictions on the correlations or shapes of the wage and location-choice error terms. The cost of such weak assumptions is that we do not estimate an intercept for each equation—it disappears in the differencing process. Without further assumptions, this method does not provide estimates of place effects.

Therefore, we impose the assumption that we have been using as a measuring stick for all the previous estimates. We assume that there are no place effects for individuals in the highest education category. We assume that the observed earnings difference of 24.7 percent (Row 1, Table 2) is entirely due to person effects. Setting predicted wages equal between the city and expanding suburbs for this group allows us to estimate place effects for the remaining individuals in the sample. Row 6 of Table 2 provides these estimates.

The estimated place effects for men with only a high school degree and with less than a high school degree are -5.4 percent and -5.8 percent, respectively. For men with some college, the place effect is -5.6 percent and for men with a bachelor's degree, the estimated effect is -2.9 percent. A 95 percent confidence interval of place effects for men with a high school degree ranges from -15.6 percent to 4.0 percent. We conclude that white men with little education cannot substantially increase their earnings by moving from the central city to the expanding suburbs. Most, if not all, of the spatial variation in earnings for white men is due to sorting by earnings and characteristics that determine earnings. Our point estimates suggest that wages are actually lower in the suburbs than in the city, as implied by a wage gradient in a monocentric city. However, the precision of our estimates does not allow us to reject the hypothesis that wages are the same in the city and the suburbs.

CONCLUSION

With few exceptions, previous work on intrametropolitan earnings variation has attributed observed differences in earnings to place effects rather than to person effects. Our work is premised on the assumption that for highly educated white men, geographic labor market barriers are sufficiently weak that all place effects will be quickly arbitraged away. Based on this assumption, we reject several variations of conventional parametric selection models that attempt to control for sorting when estimating location-specific earnings equations. Our conclusion that place effects are minimal for white men of all education levels is derived from combining our fundamental assumption regarding labor markets for highly educated white men with estimates from a selection model that places no restrictions on the functional form of the location choice equation or the distributions of the earnings and location-choice error terms.

Although we have removed some functional form assumptions that we do not believe are imposed by theory, we wish to be explicit about the remaining assumptions that are identifying the model. An earnings equation functional form assumption is necessary for identification. The arbitrary nonlinearities of the nonparametric location choice estimator act as instrumental variables that are only useful when they are excluded from the wage equation. (See Ahn and Powell, 1993, for a further discussion.)

Fortunately, the familiar log linear form for the wage equation has decades of theoretical and empirical support. Indeed, since Mincer’s (1974) justification of a linear relationship between log wages and education, experience, and experience squared, this relationship has been the mainstay of labor econometrics. Our assumption that marriage is exogenous, allowing its inclusion in the wage equation, is based on the work by Korenman and Neumark (1991). We exclude other frequently included variables such as industry, occupation, and number of children, because they are likely to affected by earnings and/or residential location, making them endogenous.

Although we have removed the functional form assumptions from the index and error distribution of the location choice equation, several assumptions remain. First, we assume that earnings enter the location utility (equation (2)) linearly with a fixed (i.e., nonrandom) coefficient γ_j. This implies that an increase in earnings will have the same impact on location choice for any individual. We also assume that the distribution of the location error term (v_{ij}) does not depend on the explanatory variables. These assumptions permit the conditional expectation of the earnings error term to be expressed as an arbitrary function of the choice probability only, rather than as an arbitrary function of all the explanatory variables.

The maintained assumption that there are no place effects for high-skilled workers is necessary for identification after assumptions regarding the distributional forms of the error terms are eliminated. Without parametric distributions, we cannot calculate the conditional expectation of the earnings error term and must treat it as a nuisance parameter. After differencing to eliminate the conditional expectation of the earnings error term, we cannot estimate the level of the earnings equations without additional assumptions. Our assumption that earnings are the same for high-skilled suburban and city residents identifies the relative levels of the two earnings equations.

The determination of whether place effects exist for low-skilled workers has important public policy ramifications. Current microeconomic policy is based on the assumption of both person and place effects. Education and job training subsidies are aimed at remediating skill deficiencies, while economic development tax credits and empowerment zones are targeted at restoring job growth to inner cities. The proper mix of labor market policies depends on the relative strength of the person and place effects. Our conclusion implies that subsidizing job-creating investments in the central city will not substantially raise earnings among inner-city residents. It should be noted, however, that there may be other benefits from such policies which our analysis of earnings does not address.

Although we have focused on white men, our conclusions have implications for central-city black workers who are the traditional subject of spatial mismatch research. If low-educated black and white workers compete for the same jobs, our findings that wages are arbitraged between the expanding suburbs and central city imply that spatial mismatch is not an explanation for the observed black/white earnings gap. In order to explain the black/white earnings gap, we must turn to other factors, such as racial discrimination in the labor market and in the processes that determine observed and unobserved skill levels.

APPENDIX
Regressions for Table 2

Regression for Row 3: Exogenous Location
Dependent Variable: Log Weekly Earnings

	City		*Expanding Suburbs*	
	Coefficient	Standard Error	Coefficient	Standard Error
Experience	.0315	.0042	.0467	.0025
$Exp^2/100$	-.0451	.0100	-.0809	.0059
Education	.0565	.0038	.0692	.0022
Married	.1779	.0215	.1911	.0150
Constant	4.540	.0736	4.362	.0408
S	.4006		.3653	
R^2	.1681		.2609	
N	1,767		4,811	

Multinomial Logit Used to Create Selection Correction Term for Row 4
City is the Omitted Category

	Expanding Suburbs			*Steel Towns and Old Suburbs*		
	Coefficient	Standard Error	z	Coefficient	Standard Error	z
Age	.1264	.0313	4.030	.0673	.0357	1.881
$Age^2/100$	-.1422	.0394	-3.610	-.0744	.0449	-1.659
Education	.8037	.0681	11.799	-.0007	.0125	-0.056
Married	.0700	.0108	6.446	.5957	.0787	7.562
Constant	-3.205	.6051	-5.297	-1.661	.6907	-2.405

Log Likelihood = -8568.1741
Number of Observations: 8,732

Regression for Row 4: Linear Location Choice and Parametric Error Distributions
Dependent Variable: Log Weekly Earnings

	City		*Expanding Suburbs*	
	Coefficient	Standard Error	Coefficient	Standard Error
Experience	.0156	.0090	.0442	.0048
$Exp^2/100$	-.0177	.0170	-.0763	.0096
Education	.0320	.0129	.0636	.0094
Married	-.0983	.1412	.1534	.0645
λ	.8391	.4240	-.1865	.3112
Constant	4.085	.2414	4.627	.4458
S	.4003		.3653	
R^2	.1699		.2610	
N	1,767		4,811	

Regression for Row 5: Nonlinear Location Choice and Parametric Error Distributions
Dependent Variable: Log Weekly Earnings

	City		*Expanding Suburbs*	
	Coefficient	Standard Error	Coefficient	Standard Error
Experience	.0324	.0044	.0429	.0026
$Exp^2/100$	-.0471	.0103	-.0732	.0061
Education	.0570	.0038	.0621	.0027
Married	.1881	.0261	.1481	.0183
λ	-.0347	.0499	-.2159	.0535
Constant	4.562	.0802	4.683	.0896
S	.4007		.3647	
R^2	.1683		.2626	
N	1,767		4,811	

Regression for Row 6: Nonparametric Selection
Dependent Variable: Log Weekly Earnings

	City		*Expanding Suburbs*	
	Coefficient	Standard Error	Coefficient	Standard Error
Experience	.0418	.0114	.0517	.0031
$Exp^2/100$	-.0845	.0273	-.0929	.0074
Education	.0536	.0073	.0713	.0022
Married	.2210	.0467	.1678	.0214
N	1,767		4,811	

Note: Standard errors for regression with nonparametric selection correction are calculated by bootstrapping.

References

Ahn, H., and J. Powell. 1993. “Semiparametric Estimation of Censored Selection Models with a Nonparametric Selection Mechanism.” *Journal of Econometrics* 58: 3–29.

Angrist, J. 1995. “Conditioning on the Probability of Selection to Control Selection Bias.” Mimeo, MIT.

Borjas, G. 1992. “Ethnicity, Neighborhoods, and Human Capital Externalities.” National Bureau of Economic Research Working Paper no. 4912.

Case, A., and L. Katz. 1991. “The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths.” National Bureau of Economic Research Working Paper no. 3705.

Engberg, J. 1996. “A Spatial Labor Market Equilibrium with Housing Market Discrimination.” Mimeo, Carnegie Mellon University.

Fan, J. 1993. “Local Linear Regression Smoothers and Their Minimax Efficiencies.” *Annals of Statistics* 21: 196–216.

Hardle, W. 1990. *Applied Nonparametric Regression.* Cambridge: Cambridge University Press.

Heckman, J. 1974. “Shadow Prices, Market Wages, and Labor Supply.” *Econometrica* 42: 679–693.

Heckman, J., and B. Honore. 1990. “The Empirical Content of the Roy Model.” *Econometrica* 58 (September): 1121–49.

Heckman, J., and V. J. Hotz. 1989. “Choosing among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training.” *Journal of the American Statistical Association* 84: 862–874.

Heckman, J., H. Ichimura, J. Smith, and P. Todd. 1995. “Non-Parametric Characterization of Selection Bias Using Experimental Data: A Study of Adult Males in JTPA.” Mimeo, University of Chicago.

Holzer, H. 1991. “The Spatial Mismatch Hypothesis: What Has the Evidence Shown?” *Urban Studies* 28: 105–122.

Ihlanfeldt, K. 1988. “Intra-Metropolitan Variation in Earnings and Labor Market Discrimination: An Econometric Analysis of the Atlanta Labor Market.” *Southern Economic Journal* 50: 123–140.

Ihlanfeldt, K. 1992. “Intraurban Wage Gradients: Evidence by Race, Gender, Occupational Class, and Sector.” *Journal of Urban Economics* 32: 70–91.

Ihlanfeldt, K. 1994. “The Spatial Mismatch between Jobs and Residential Locations within Urban Areas.” *Cityscape* 1: 219–244.

Ihlanfeldt, K., and D. Sjoquist. 1990. “Job Accessibility and Racial Differences in Youth Employment Rates.” *American Economic Review* 80: 267–276.

Jencks, C., and S. Mayer. 1990. “Residential Segregation, Job Proximity, and Black Job Opportunities.” In *Inner-City Poverty in the United States,* edited by L. Lynn, Jr., and M. McGeary. Washington, D.C.: National Academy Press.

Kain, J. 1968. “Housing Segregation, Negro Employment, and Metropolitan Decentralization.” *Quarterly Journal of Economics* 2: 175–197.

Kain, J. 1992. “The Spatial Mismatch Hypothesis: Three Decades Later.” *Housing Policy Debate* 3.

Korenman, S., and D. Neumark. 1991. “Does Marriage Really Make Men More Productive?” *Journal of Human Resources* 26: 282–307.

Lee, L-F. 1983. “Generalized Econometric Models with Selectivity.” *Econometrica* 51: 507–512.

Manski, C. 1993. “Identification Problems in the Social Sciences.” *Sociological Methodology* 23: 1–56.

Manski, C., G. Sandefur, S. McLanahan, and D. Powers. 1992. “Alternative Estimates of the Effect of Family Structure during Adolescence on High School Graduation.” *Journal of the American Statistical Association* 87: 25–37.

McMillen, D. 1993. “Can Blacks Earn More in the Suburbs? Racial Differences in Intra-metropolitan Earnings Variation.” *Journal of Urban Economics* 33: 135–150.

Mills, E., and B. Hamilton. 1994. *Urban Economics,* 5th ed. New York: Harper Collins.

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

Mroz, T. 1987. “The Sensitivity of an Empirical Model of Married Women’s Hours of Work to Economic and Statistical Assumptions.” *Econometrica* 55: 765–799.

Newey, W., J. Powell, and J. Walker. 1990. “Semiparametric Estimation of Selection Models: Some Empirical Results.” *American Economic Review Papers and Proceedings* 80: 324–328.

O’Regan, K., and J. Quigley. 1995. “Spatial Contributions to Employment Outcomes: The Case of New Jersey Teenagers.” Mimeo, Yale University.

Raphael, S. 1995. “The Spatial Mismatch Hypothesis of Black Youth Unemployment: Evidence from the San Francisco Bay Area.” Mimeo, University of California at Berkeley.

Robinson, C. 1989. “The Joint Determination of Union Status and Union Wage Effects: Some Tests of Alternative Models.” *Journal of Political Economy* 97: 639–667.

Rosenbaum, P. R., and D. B. Rubin. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika* 70: 41–55.

Roy, A. D. 1951. “Some Thoughts on the Distribution of Earnings.” Oxford Economic Papers 3: 135–146.

Strazheim, M. 1980. “Discrimination and the Spatial Characteristics of the Urban Labor Market for Black Workers.” *Journal of Urban Economics* 7: 119–140.

Tiebout, C. 1956. “A Pure Theory of Local Expenditures.” *Journal of Political Economy* 64: 416–424.