Social Interactions, Thresholds, and Unemployment in Neighborhoods *

Brian Krauth
Department of Economics
Simon Fraser University
Burnaby BC V5A 1S6
mailto:bkrauth@sfu.ca
http://www.sfu.ca/~bkrauth

November 8, 1999

Abstract

This paper finds that the predicted unemployment rate in a community increases dramatically when the fraction of neighborhood residents with college degrees drops below twenty percent. This threshold behavior provides empirical support for "epidemic" theories of inner-city unemployment. Using a structural model with unobserved neighborhood heterogeneity in productivity due to sorting, I show that sorting alone cannot generate the observed thresholds without also implying an implausible shape for the wage distribution. This provides further evidence that true social interaction effects are driving the earlier results.

*I have received helpful commentary from William A. Brock, Kim-Sau Chung, Steven Durlauf, Giorgio Fagiolo, Peter Norman, Jonathan Parker, and Krishna Pendakur, as well as seminar participants at Simon Fraser University, SUNY at Albany, University of California - San Diego, University of Wisconsin, and the NBER Summer Workshop. An earlier draft of this paper was called "Social Interactions and Aggregate Neighborhood Outcomes." Revisions of this paper are available at the address above.

1 Introduction

The work of sociologist William Julius Wilson [21, 22] has inspired social scientists to investigate the role of neighborhoods in perpetuating and exacerbating social and economic problems among the poor. In *The Truly Disadvantaged*, Wilson argues for a "concentration effects" or "social isolation" explanation for the social pathologies and high unemployment rates found in Chicago's poor and predominantly African-American neighborhoods. Without the stabilizing social influence and connections to the larger society of steadily employed middle class neighbors, residents of a community in which the economically vulnerable are concentrated find it more difficult to get and hold jobs. Social interaction effects thus reinforce the direct effects of structural change in labor markets and the departure of middle class families. Wilson claims that this self-reinforcing mechanism can explain the explosion in unemployment and other social problems seen in Chicago's inner-city communities since 1970. Partly in response to Wilson's work, a number of economists have investigated the empirical relevance of social interaction effects, as well as their aggregate or equilibrium implications.

The possibility of threshold effects in neighborhoods, or what Crane [3] calls the "epidemic theory of ghettos," has been addressed by a number of theoretical papers in economics [2, 6, 14, 19]. These models show that social interaction effects can lead to the existence of multiple equilibria in neighborhoods. In a dynamic context, these models imply individual neighborhoods can experience "epidemics" or "tipping", i.e., discontinuous changes over time as the neighborhood moves between equilibria. In a cross-sectional context these same models can imply thresholds, or a discontinuous relationship between neighborhood resources and neighborhood outcomes. Epidemics, tipping, and thresholds are of interest both in explaining the observed dynamics of neighborhoods and other social groups, and in constructing public policy. Sudden large changes in the prevalence of social behavior, in the absence of corresponding large changes in individual characteristics, often become areas of active public and academic interest. One example is Wilson's argument described above. Another is the anti-crime policy of the Giuliani administration in New York City, which has explicitly applied James Q. Wilson and George Kelling's [12] "broken windows" theory of crime. This theory, which suggests that small breakdowns in neighborhood order lead to both more disorder and more serious criminal activity, is closely related to epidemic models of social interactions. This shift in policy has coincided with a substantial drop in New York's crime rate. If this outcome can be interpreted as an example of a threshold effect induced by social interactions, it represents significant opportunities for public policy, as the positive results seen in some communities can be transferred to others. By the same token, a true threshold effect in unemployment also represents a significant opportunity for public policy. The empirical relevance of theoretical models with neighborhood unemployment thresholds is thus of substantial policy interest.

However, few empirical papers have addressed the possibility of neighborhood thresholds at all. Instead, the empirical literature has focused almost exclusively on the existence of neighborhood effects in individual level data. In this paper, I use census tract data from twenty large United States cities to determine whether there exists evidence for a threshold in the neighborhood-level relationship between unemployment and neighborhood human capital. Using an empirical approach which is similar in spirit to that of Crane [3], I find that in almost all cities, unemployment increases dramatically when the percentage of residents with college (Bachelor's) degrees falls below a critical value near twenty percent. This stylized fact suggests that thresholds induced by social interactions are a characteristic of neighborhood employment dynamics.

Unlike much of the empirical literature on social interactions, I explicitly consider alternative explanations for this stylized fact. Positive sorting, or the tendency of individuals to live near others who are similar, can produce spurious signs of social interactions and neighborhood thresholds.[1] To address the issue of sorting, I develop a parametric model which explicitly incorporates neighborhood formation. I show that, if sorting-induced correlation between neighbors in unobserved productivity is to explain the thresholds observed in the data, this unobserved productivity must have a distribution which is implausible for a productivity variable. In contrast, a model with social interactions can generate the observed threshold, providing additional evidence that this threshold is not spurious.

[1] Bénabou [1] and Durlauf [4] show that positive sorting is in fact a likely outcome of neighborhood formation by optimizing agents when there are social interactions.

2 Do neighborhoods have thresholds?

In this section I estimate a nonparametric regression $E(\bar{y}_n|\bar{x}_n)$ of a neighborhood's unemployment rate ($\bar{y}_n$) on its average human capital ($\bar{x}_n$). Each neighborhood's average human capital is measured as the fraction of residents who have a Bachelor's degree. The primary question of interest is whether this regression shows a discontinuity or threshold.

2.1 Data

The data source for this study is the 1990 Census Summary Tape File 3A. Each neighborhood observation is a census tract from one of twenty selected cities in the United States.[2] A census tract is a geographic region with population usually between 2,500 and 8,000 residents which provides a rough approximation to neighborhood. Tracts are omitted if they contain fewer than 300 residents in the labor force. Tracts are initially drawn by the Census Bureau to be relatively homogeneous with respect to economic and demographic characteristics. However, the boundaries of many urban tracts were drawn several decades before 1990, so they may imperfectly reflect current neighborhoods. In addition, geography may not accurately reflect social contact. Even in mixed neighborhoods, there may not be much social contact between members of different ethnic, educational, or income groups. As a result, using census tracts to represent neighborhoods may understate the true importance of social interactions.

Neighborhood unemployment rates, which will be denoted by $\bar{y}_n$, are calculated for male and female civilians in the labor force. Neighborhood human capital, denoted by $\bar{x}_n$, is measured as the fraction of neighborhood residents over age 25 who have Bachelor's degrees. I estimate the regressions separately for each of the twenty cities in order to control for variation across labor markets. Table 1 shows the mean and median values of the two variables, as well as the number of tracts in each city. As the table shows, different cities experience different outcomes, supporting the treatment of each city as a separate labor market. The cities in the sample represent the five largest Metropolitan Statistical Areas (MSA) or Consolidated Metropolitan

[2] Topa [20] also studies unemployment in urban census tracts. In contrast to this paper, Topa's focus is on the patterns of correlation in unemployment between neighboring tracts rather than patterns within tracts.

Statistical Areas (CMSA) in each of four regions - Northeast, Midwest, South, and West. An MSA is a primarily urban county or collection of several contiguous counties. Each MSA is constructed by the Census Bureau to roughly represent a single labor market. A CMSA is a collection of contiguous metropolitan areas each of which is called a Primary Metropolitan Statistical Area (PMSA). I use CMSA rather than PMSA or incorporated city in this paper for several reasons. The political boundaries of most major cities are generally much narrower than the corresponding metropolitan labor market, as evidenced by substantial commuting between city and suburbs. PMSA boundaries, while wider than city boundaries, tend to be inconsistent across different metropolitan areas. For example, the Los Angeles - Long Beach PMSA contains only Los Angeles County, while the Houston PMSA contains eight different counties. A CMSA can include sparsely populated areas where a census tract is a poor proxy for a neighborhood. However, there are few such tracts relative to the number of urbanized tracts, so the influence of these tracts is slight. Estimation at the PMSA or county level (not reported) does not produce substantially different results from those found using the entire CMSA.

2.2 Econometric issues

2.2.1 Identification issues

The empirical relevance of neighborhood social interaction effects is a matter of substantial disagreement among social scientists. Manski [15] argues that this disagreement results from a fundamental problem of identification. For example, we observe that a child is more likely to smoke cigarettes if his friends smoke. This observation can be explained by social interaction effects - he smokes because his friends smoke. It can also be explained by "sorting" - he smokes because he likes to smoke, and he has made friends with fellow smokers. The identification problem is that there is no way to tell the difference. Attempts to infer social interaction effects from behavior are subject to this "sorting critique" whenever the neighborhood is selected by purposive economic agents rather than by random experiment.

Social interaction effects and sorting have very different policy implications. If there are true social interaction effects in unemployment, the residential distribution of individuals has an impact on individual employment outcomes. Policies such as the Chicago Housing Author-

City	# of Tracts	% w/Bachelors Mean	% w/Bachelors Median	Unemployment % Mean	Unemployment % Median
Baltimore	566	21.1	18.0	5.8	3.8
Boston	1158	25.7	22.1	7.6	6.3
New York	4813	23.4	19.5	7.6	5.7
Philadelphia	1424	21.9	17.4	6.1	4.2
Washington, D.C.	887	36.3	35.0	4.4	3.1
Chicago	1841	20.3	14.4	9.8	5.9
Cleveland	809	16.7	11.9	9.0	5.4
Detroit	1239	17.6	12.4	10.5	6.6
Minneapolis - St. Paul	612	25.2	21.5	5.3	4.3
St. Louis	443	18.5	13.6	8.3	5.7
Atlanta	469	24.4	20.1	6.2	4.5
Dallas - Ft. Worth	816	24.8	21.3	6.7	5.2
Houston	780	20.7	14.7	7.7	6.3
Miami - Ft. Lauderdale	424	18.1	14.1	7.5	6.6
Tampa - St. Petersburg	400	16.6	13.8	5.8	4.8
Los Angeles	2504	21.8	18.3	6.9	5.9
Phoenix	455	21.7	20.1	6.4	5.4
San Diego	421	24.1	20.9	6.3	5.5
San Francisco - Oakland	1281	30.4	27.8	5.5	4.4
Seattle	511	26.1	22.7	4.9	4.1

Table 1: Summary statistics for the 1990 Census, by city.

ity's attempts to relocate public housing tenants into more economically diverse neighborhoods [17] are based in part on the belief that the residential concentration of the very poor produces a socially isolated "underclass" with little hope of better economic outcomes. The effectiveness of such a policy depends critically on the existence of economically significant social interaction effects. As a result, distinguishing between social interaction effects and sorting effects is important for the development of social policy.

One approach to identification is to ask individuals directly how they make choices. In many cases, survey data on how individuals make choices provides direct support for some degree of peer influence. For the case of employment, survey data indicate that around half of workers find their jobs through referrals from employed friends, family, and neighbors [8]. However, the fact that individuals often use social resources in job finding does not say much about their employment prospects in the absence of these resources.

In order to address the question of economic importance, one must compare the experience of observationally similar people in different social environments. In most of the empirical literature on neighborhood effects, this takes the form of estimating a probit or logit model on individual data linked with information on the person's neighborhood:

$$\Pr(y_i = 1 | x_i, \bar{x}_n) = \Lambda\left(\beta_0 + \beta_1 x_i + \beta_2 \bar{x}_n\right) \tag{1}$$

where y_i is the outcome, x_i is a vector of individual characteristics, and $\bar{x}_n$ is some variable describing neighborhood composition. It is tempting to say that the coefficient β_2 measures the social interaction effect, but it actually measures the combined social interaction effect and sorting effect. The identification problem is that additional assumptions are needed to estimate the two effects separately.

The most common assumption is exogenous selection into neighborhoods. Exogenous selection means that any individual characteristics which affect both neighborhood and the outcome in question are included in the regressors. In this case, there is no sorting effect (though there may be sorting), and β_2 is the social interaction effect. Unfortunately, the assumption of exogenous selection is difficult to justify in most economically interesting cases. It would be surprising if school dropout, criminal behavior, public aid receipt, or out-of-wedlock childbearing did not affect neighborhood choice, yet those are among the primary individual outcomes analyzed in this literature. If the outcome itself affects neighborhood choice, any variable affecting the outcome also affects neighborhood. For example, employment status has a direct impact on income, and income has a clear impact on choice of residential location. As a result, the assumption of exogenous selection cannot be used for the case of unemployment.

Several recent articles [5, 7] dispense with the unpalatable assumption of exogenous selection and instead use instrumental variables to adjust for any endogeneity. Others [18] use the results from natural experiments. However, usable instruments or natural experiments are rare. In this paper, I follow a new approach by developing a parametric model of the sorting process. Some parameters of this model can be estimated under the hypothesis of no social interactions. If these implied parameter values contradict known results from other studies, the assumption of no social interactions can be rejected. The details and results of this model are outlined in Section 3.

2.2.2 Endogenous versus contextual effects

Manski [15] also distinguishes between "endogenous" and "contextual" social interaction effects. Effects are endogenous if individual decisions are affected directly by the decisions of others. Effects are contextual if individual decisions are affected directly by the background characteristics of others. For example, contact with employed neighbors may increase one's own probability of employment. Contact with educated neighbors may do so as well. A positive impact of neighbors' employment on own employment is an endogenous effect, while a positive impact of neighbors' education level on own employment is a contextual effect. As with distinguishing between social interaction effects and sorting effects, distinguishing between endogenous and contextual effects faces an identification problem. In Section 3, I use restrictions on the form of the two effects to test for the presence of economically significant endogenous effects.

The relevance of the distinction between endogenous and contextual effects lies in the fact that endogenous effects imply a social multiplier. A policy intervention which directly assists some residents of a neighborhood in finding employment will indirectly help the other residents. As a result, broad-based policies will have larger impacts on group outcomes than the same policy will on an isolated individual. If endogenous effects are strong enough, multiple equilibria, thresholds, and epidemics or tipping phenomena are generated. In contrast, contextual effects do not imply a social multiplier.

2.2.3 Aggregation

Aggregate data on neighborhoods have many useful properties - most notably, the availability of large numbers of aggregates and the fact that the Census includes the entire population of neighborhoods in a city. Both of these features are exploited in this study to achieve very high precision in estimates and to model the neighborhood formation process itself. Unfortunately, it is not possible in general to infer individual-level parameters from aggregates.

The standard method for ascertaining the existence of social interaction effects is to estimate Equation (1) using individual-level data and test the null hypothesis $\beta_2 = 0$. A nonparametric analogue would be to estimate:

$$E(y_i|x_i, \bar{x}_n) \tag{2}$$

The joint null hypothesis of exogenous selection and no social interaction effects is equivalent to:

$$H_0 : E(y_i|x_i, \bar{x}_n = X) - E(y_i|x_i, \bar{x}_n = X') = 0 \quad \text{for all } \{X, X'\} \tag{3}$$

This null hypothesis is easy to test using individual data, but can it be tested using only neighborhood-level averages $\bar{y}_n$ and $\bar{x}_n$? In general, no. However, if x_i is a single binary variable, then:

$$\begin{aligned} E(\bar{y}_n|\bar{x}_n = X) &= E\left(y_i|x_i = 0, \bar{x}_n = X\right) \qquad (4) \\ &+ \bar{x}_n \left[E\left(y_i|x_i = 1, \bar{x}_n = X\right) - E\left(y_i|x_i = 0, \bar{x}_n = X\right)\right] \end{aligned}$$

Under the null hypothesis (3), Equation (4) is linear in $\bar{x}_n$. In other words, as long as x_i is a single binary variable, testing for the linearity of $E(\bar{y}_n|\bar{x}_n)$ is analogous to testing $\beta_2 = 0$ with individual data, the key exercise in much of the previous literature.[3]

While any form of social interaction effect is of interest, threshold effects are of particular interest. A threshold nonlinearity in principle is simply a discontinuity in the regression function. However, it is difficult in practice to empirically distinguish a discontinuous regression function (Figure 1) from one that has a steep slope over a short range of the dependent variable (Figure 2). As a result, any continuous regression function that shows a large change in unemployment over a small range of neighborhood average human capital can be said to

[3] If $E(y_i|x_i, \bar{x}_n)$ is linear and strictly monotonic in $\bar{x}_n$, then $E(\bar{y}_n|\bar{x}_n)$ will also be linear. The aggregate test will fail to reject the null hypothesis (3) , even though it is false and would be rejected by the individual-level test.

provide evidence for neighborhood thresholds. No formal criteria for whether a particular change is "large" will be defined; instead that judgment is left to the reader.

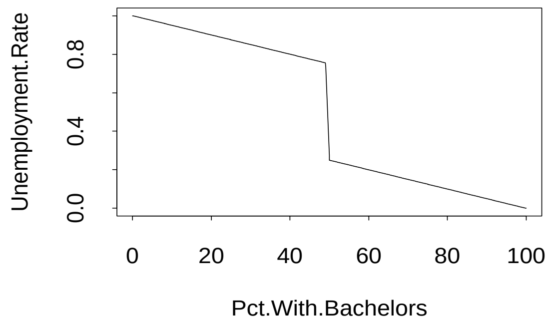

Figure 1: An example of a threshold relationship between two variables.

2.3 Baseline results

Figure 3 shows a scatter plot of neighborhood unemployment ($\bar{y}_n$) versus neighborhood educational attainment ($\bar{x}_n$) for the Chicago CMSA. The figure shows an interesting set of patterns which appear in the other cities as well. For neighborhoods with more than twenty percent college graduates, unemployment rates are uniformly low. In contrast, neighborhoods with fewer than twenty percent college graduates appear to have much higher average unemployment rates as well as much higher variability in unemployment.

Figures 4 through 7 show nonparametric regressions of neighborhood unemployment on neighborhood human capital for each of the cities in the sample. These estimates are calculated using the supersmoother [9, page 181], and the 95 percent confidence intervals shown are estimated by the bootstrap with 1,000 iterations.[4] As appears in

[4]To describe the procedure in more detail, I estimate a series of $X\%$ pointwise intervals, then increase the value of X until 95% of the bootstrapped regression functions are entirely within these intervals (see Härdle [9] for a discussion of this procedure). As a result, the confidence interval has a 95% probability of containing the *entire* true regression function.

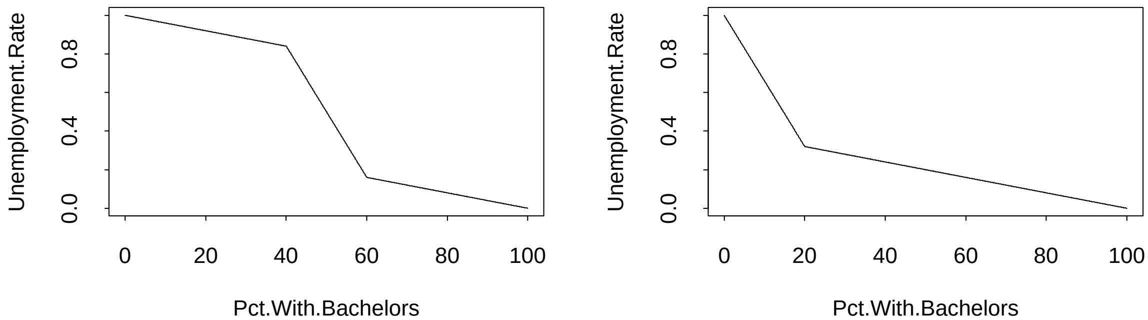

Figure 2: Examples of regression relationships which provide evidence for thresholds.

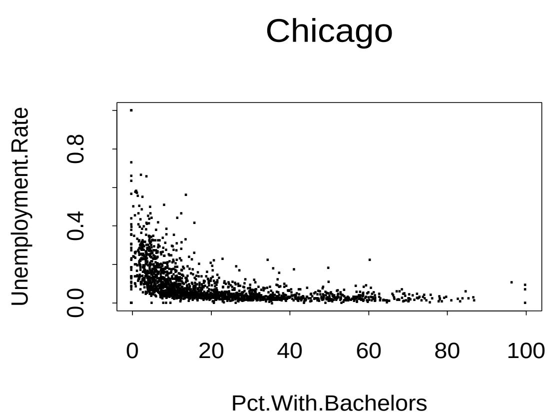

Figure 3: Scatter plot of neighborhood unemployment versus neighborhood education in Chicago CMSA.

the scatter plot, the regression relationship is noticeably nonlinear, and almost every city exhibits a clear threshold. The predicted unemployment rate increases substantially when the percentage of college graduates in a neighborhood falls below about twenty. This threshold is consistent with epidemic models of social interactions.

The apparent nonlinearity of the regressions can be placed in a more formal statistical setting. I test for the linearity of each regression using Härdle and Mammen's nonparametric method [10] for testing the linearity of a conditional expectation function. The procedure is to estimate the CEF both nonparametrically and with OLS. The average Euclidean distance between the two estimators at each point of the support, with a few bias corrections described in their paper, is then the test statistic. Härdle and Mammen show that this test statistic is consistent and asymptotically normal, but show that a bootstrap estimator yields better small-sample estimates for the distribution of the test statistic under the null hypothesis of linearity. Table 2 shows the estimated test statistic, critical value, and p-value for this linearity test by city. The test statistic is calculated as in Härdle and Mammen, and its distribution under the null is estimated using the wild bootstrap with 1,000 iterations. As the table shows, linearity of $E(\bar{y}_n|\bar{x}_n)$ is easily rejected by the data in all cases. As shown in Section 2.2.3, this also leads us to reject the related joint hypothesis of exogenous selection and no social interactions. At minimum, there is some sorting effect or social interaction effect.

While linearity is easily placed within a formal hypothesis test setting, I have chosen to define "threshold" informally as a large change in average outcome associated with a small change in the regressor. Establishing whether the true conditional expectation function for a given city is characterized by thresholds cannot be done formally. However, there is still convincing evidence that this is the case. The thresholds appear across the many different cities in this sample, and the confidence intervals for Figures 4 - 7 are quite narrow due to the large samples. This is strong statistical evidence that the threshold shapes are characteristics of the underlying CEF.

Thresholds are a robust characteristic of the reduced form relationship between neighborhood resources and neighborhood unemployment. This reduced-form result is consistent with epidemic models of social interaction effects. However, other structural models may have similar reduced-form implications. In Section 3, I consider alternative explanations and argue that the epidemic explanation is more

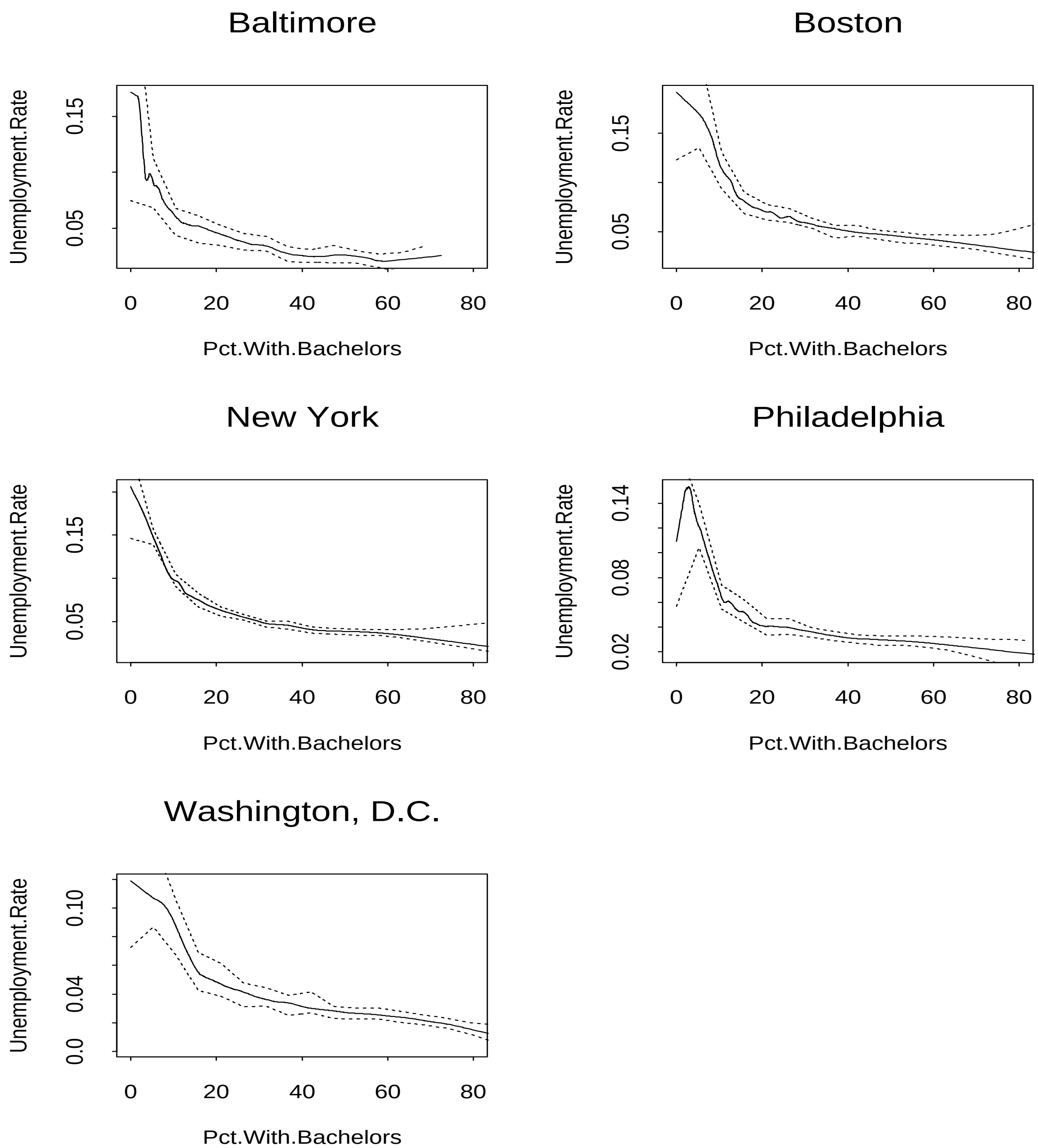

Figure 4: Nonparametric regressions for northeastern cities.

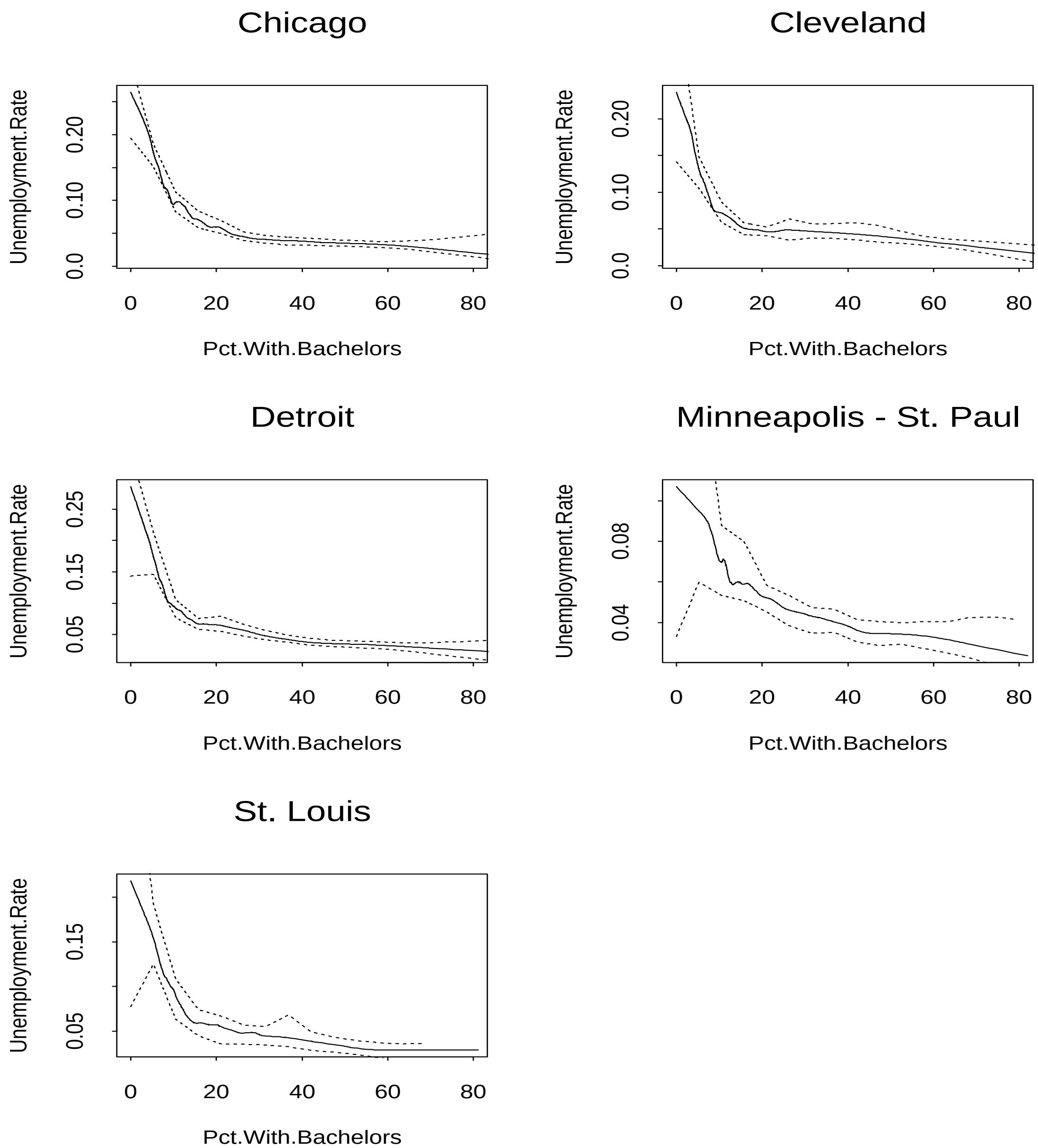

Figure 5: Nonparametric regressions for midwestern cities.

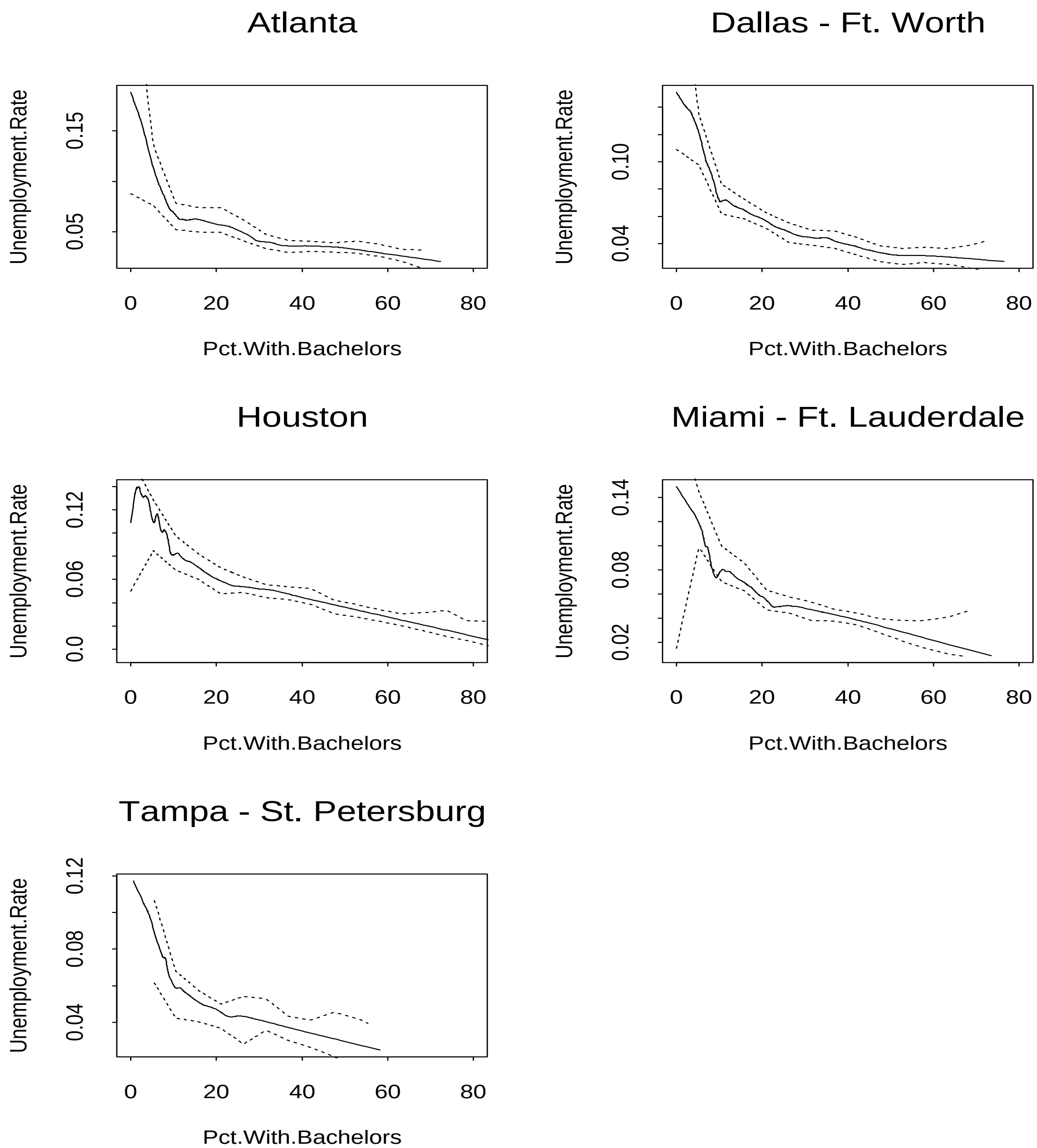

Figure 6: Nonparametric regressions for southern cities.

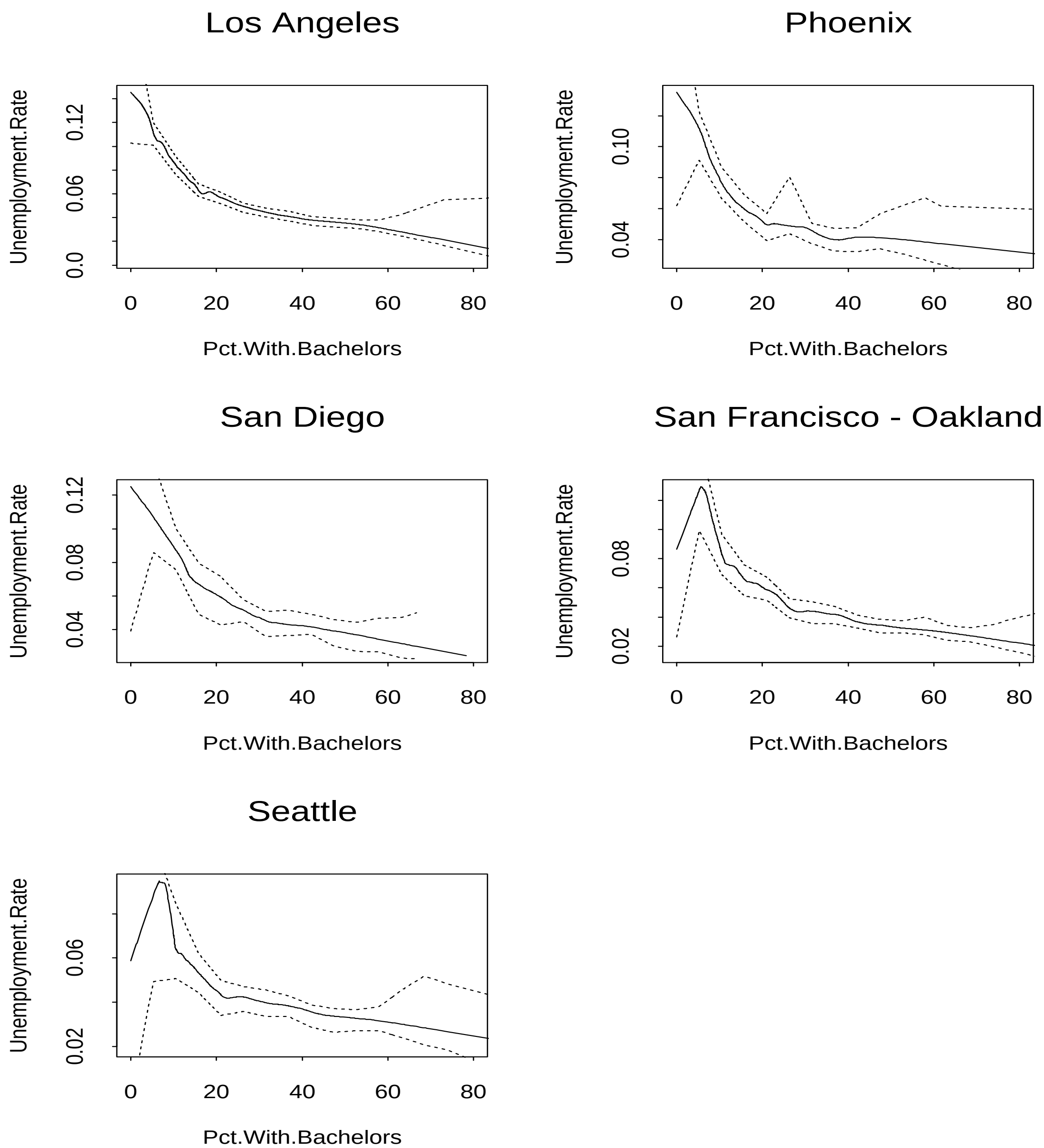

Figure 7: Nonparametric regressions for western cities.

plausible.

City	Test Statistic	95% Critical Value	P-value
Baltimore	0.338	0.0139	0
Boston	0.359	0.00512	0
New York	1.76	0.000696	0
Philadelphia	1.02	0.0101	0
Washington, D.C.	0.337	0.0137	0
Chicago	3.45	0.00449	0
Cleveland	1.52	0.0207	0
Detroit	1.59	0.00242	0
Minneapolis - St. Paul	0.862	0.00924	0
St. Louis	0.356	0.0144	0
Atlanta	0.108	0.00236	0
Dallas - Ft. Worth	0.346	0.00214	0
Houston	0.0586	0.000758	0
Miami - Ft. Lauderdale	0.0989	0.00598	0
Tampa - St. Petersburg	0.083	0.0221	0
Los Angeles	0.444	0.000594	0
Phoenix	0.476	0.0117	0
San Diego	0.00191	0.00064	0
San Francisco - Oakland	0.245	0.00139	0
Seattle	0.0572	0.00862	0

Table 2: Results from Härdle-Mammen test for linearity of $E(\bar{y}_n|\bar{x}_n)$

3 Alternative explanations, sorting, and misspecification

In general, separate identification of sorting effects and true social interaction effects cannot be made from observations on behavior without strong identifying restrictions on the nature of the effects. The previous section established that either sorting effects or social interaction effects must be present. In this section, I apply a new strategy to distinguish between the two types of effects.

3.1 How do individuals sort?

The residential location choice made by individuals and families is influenced by many family characteristics – current location, income, taste for housing quality, family size, presence of family or social ties, ethnicity, and many others. However, only those factors which can lead to mistaken inference of social interactions are of interest for this paper. In order to do so, a candidate sorting variable must be correlated with neighborhood educational level and also help to improve predictions of employment probability after controlling for individual education level. For example, sorting on educational attainment cannot produce spurious thresholds, since this is the independent variable. Another alternative is that individuals sort on employment status itself. Taken literally, this implies that the distribution of neighborhood unemployment rates should have many neighborhoods with no unemployment and a few neighborhoods with unemployment of 100 percent. Yet nearly all of the neighborhoods in my sample have unemployment rates which strictly between zero and thirty percent. Direct sorting on employment status can thus be ruled out.

A more promising candidate is that families sort on income. In a simple housing market with no externalities, neighborhood stratification on income level will occur if the most attractive and expensive housing locations are in the same neighborhood. As housing is a normal good, high-income families will choose to live in the most attractive neighborhood.[5] However, they are unlikely to sort exclusively on current income. More than half of the sample lived in the same location in 1985 and 1990, and even the highest-education neighborhoods had some unemployment. These two facts suggest that families do not always change residential location in response to a transitory change in income or employment. Indeed, if families face moving costs, or if they simply wish to smooth their consumption of housing and neighborhood resources, they choose residential location on the basis of more long-term income prospects. In the remainder of Section 3, I use an indirect approach to exploit the idea that the primary

[5] In the presence of quantitatively important social interaction effects, the neighborhood formation problem becomes more complex because families care who their neighbors are. Bénabou [1] and Durlauf [4] describe conditions under which social interaction effects reinforce the incentive to sort.

determinant of neighborhood choice is a family's long-term income prospects, which I refer to as their "productivity." Note that productivity should not be interpreted as innate ability, IQ, education, or a test score. Instead, productivity includes any characteristics of the worker which affect the worker's income-generating ability which are permanent and portable.

3.2 A simple model

In this section, I develop a model in which each individual's employment outcome is described by a simple binary choice model and individuals sort into neighborhoods on an unobserved productivity variable. Even with these assumptions separate identification of sorting and social interaction effects is not possible. Instead, I assume that there is no social interaction effect, estimate the model in which sorting must explain the threshold, and evaluate the resulting estimates for plausibility.

In the model, each of the I workers in a given city is characterized by education x_i and productivity $z_i \in [0,1]$, both of which are exogenous. As described above, a worker's productivity level is his or her long-term income-generating ability, measured in dollars. Education level is binary:

$$x_i = \begin{cases} 0 & \text{if no college degree} \\ 1 & \text{if college degree} \end{cases}$$

These two characteristics are distributed jointly across individuals with distribution function F. The conditional distribution $F(z_i|x_i)$ has a strictly monotone likelihood ratio, i.e., an individual with a college degree is more likely to have higher productivity. Each worker goes through a two-step process:

1. Worker chooses a neighborhood based on a simple sorting rule. Workers stratify perfectly into N neighborhoods of equal size based on their value of z_i. In other words, neighborhood one contains the individuals with the I/N lowest realizations of z_i, neighborhood two contains the I/N next lowest, etc. For simplicity, assume that I/N is an integer and that any ties are broken by lottery.

2. Worker receives a random employment offer and accepts or re-

jects it, producing the binary employment outcome y_i, where

$$y_i = \begin{cases} 0 & \text{if accept (employed)} \\ 1 & \text{if reject (unemployed)} \end{cases}$$

Let $\bar{x}_n$ represent the neighborhood average of x_i, and define $\bar{y}_n$ and $\bar{z}_n$ similarly. The net wage offer (wage offer minus reservation wage) is linear in education, productivity, and any neighborhood variables:

$$w_i = \beta_0 + \beta_1 x_i + \beta_2 \bar{x}_n - \beta_3 \bar{y}_n + \beta_4 z_i + \bar{\epsilon}_n + \epsilon_i \tag{5}$$

The coefficient on $\bar{x}_n$ thus corresponds to the contextual neighborhood effect, and the coefficient on $\bar{y}_n$ corresponds to the endogenous neighborhood effect. As workers sort into neighborhoods on z_i, its coefficient corresponds to the sorting effect.

The worker accepts the net wage offer w_i if it is positive, or if the gross wage offer exceeds his or her reservation wage.

$$y_i = \begin{cases} 0 & \text{if } w_i \geq 0 \\ 1 & \text{if } w_i < 0 \end{cases} \tag{6}$$

The variables $\bar{\epsilon}_n$ and ϵ_i represent unobserved neighborhood and individual level shocks, respectively. Unlike the unobserved productivity variable z_i, which is known by the worker before choosing location, the shocks $\bar{\epsilon}_n$ and ϵ_i affect the worker after the locational choice has been made. I assume the neighborhood level shock is mean-zero conditional on the other neighborhood characteristics, though it may be heteroscedastic:

$$E(\bar{\epsilon}_n | x_i, z_i, \epsilon_i \quad \forall i \text{ in neighborhood}) = 0 \tag{7}$$

I also assume that ϵ_i is independent and identically distributed across individuals, is independent of individual characteristics, and has a logistic distribution. Unemployment thus follows a logit model:

$$\Pr(y_i = 1 | x_i, z_i, \bar{x}_n, \bar{y}_n, \bar{\epsilon}_n) = \Lambda\left(-(\beta_0 + \beta_1 x_i + \beta_2 \bar{x}_n - \beta_3 \bar{y}_n + \beta_4 z_i + \bar{\epsilon}_n)\right) \tag{8}$$

where:

$$\Lambda(X) \equiv \Pr(\epsilon_i \leq X) = \frac{e^X}{1 + e^X}$$

Note that equation (8) includes the individual-level variables x_i and z_i, while the data consist of neighborhood-level proportions. The assumptions of the sorting model allow equation (8) to be rewritten entirely in terms of aggregates.

For a large economy, the sorting model implies a strictly monotonic function $z(\bar{x}_n)$ such that every worker in a neighborhood with a fraction $\bar{x}_n$ of college graduates has productivity level $z_i = z(\bar{x}_n)$. Equation (8) becomes:

$$\Pr(y_i = 1|x_i, z_i, \bar{x}_n, \bar{y}_n, \bar{\epsilon}_n) = \Lambda\left(-(\beta_0 + \beta_1 x_i + \beta_2 \bar{x}_n - \beta_3 \bar{y}_n + \beta_4 z(\bar{x}_n) + \bar{\epsilon}_n)\right) \tag{9}$$

It is thus possible in principle to distinguish between z_i, $\bar{\epsilon}_n$, and ϵ_i, because z_i has a functional relationship with $\bar{x}_n$, but $\bar{\epsilon}_n$ and ϵ_i are uncorrelated with $\bar{x}_n$. The source of distinction is that z_i affects residential choice and $\bar{\epsilon}_n + \epsilon_i$ does not.

If the neighborhood is large, $\bar{y}_n$ converges to its expected value, producing the following aggregate structural model:

$$\begin{aligned} \bar{y}_n &= \bar{x}_n * \Lambda\left(-(\beta_0 + \beta_1 + \beta_2 \bar{x}_n - \beta_3 \bar{y}_n + \beta_4 z(\bar{x}_n) + \bar{\epsilon}_n)\right) \\ &+ (1 - \bar{x}_n) * \Lambda\left(-(\beta_0 + \beta_2 \bar{x}_n + \beta_3 \bar{y}_n + \beta_4 z(\bar{x}_n) + \bar{\epsilon}_n)\right) \end{aligned} \tag{10}$$

The approximation that $\bar{y}_n = E(\bar{y}_n)$ is reasonable for a census tract as the standard deviation of an average of a few thousand independent and identically distributed Bernoulli random variables is quite small. For example, in the median tract in Chicago, with about 2400 adults in the labor force and 6% unemployment, the standard deviation is less than one half of one percentage point.

Equation (10) will be used as the basis for estimation in this section. The parameters of interest are β_2 (the contextual effect) and β_3 (the endogenous effect), as well as the function $\beta_4 z(\bar{x}_n)$ (the sorting effect). Because both z_i and $\bar{\epsilon}_n$ are unobserved, substantial further identifying restrictions on the data-generating process would be needed to estimate the relevant parameters. Instead, I follow a more indirect strategy. I isolate each of the three effects (contextual, endogenous, and sorting) in turn by assuming the other two effects are not present. I then estimate the model under that restriction, and evaluate how well the resulting model fits the joint distribution of $\bar{y}_n$ and $\bar{x}_n$. Although it is likely that all three effects are present to some extent, this procedure will identify which effects *must* be present in order to fit the data. In Section 3.3, I evaluate how sorting explains the data in the absence of true social interactions. In Section 3.4, I compare how well contextual and endogenous effects explain the data in the absence of sorting. In the interests of space, I show results for the Chicago CMSA only. Chicago has been studied at the neighborhood level more than any other U.S. city [20, 21, 22], and the results in other cities (available from the author) are quite similar.

3.3 Can sorting generate thresholds?

In this section I assume that there are no social interaction effects, and evaluate whether sorting alone can plausibly generate the thresholds seen in the data. I find that sorting can generate thresholds, but only under an implausible distribution for the productivity variable.

Assume no social interaction effects ($\beta_2 = \beta_3 = 0$), and fix the direct benefit of education at zero ($\beta_1 = 0$).[6] Under these restrictions, equation (10) reduces to the following:

$$\bar{y}_n = \Lambda\left(-(\beta_0 + \beta_4 z\left(\bar{x}_n\right) + \bar{\epsilon}_n)\right) \tag{11}$$

A particularly convenient transformation is to take the log odds ratio of both sides. The log odds ratio function is simply the inverse of the logistic function:

$$\Lambda^{-1}\left(X\right) = \ln(X) - \ln(1 - X) \tag{12}$$

Applying the transformation to (11):

$$-\Lambda^{-1}\left(\bar{y}_n\right) = \beta_0 + \beta_4 z\left(\bar{x}_n\right) + \bar{\epsilon}_n \tag{13}$$

Taking expectations and rearranging:

$$z\left(\bar{x}_n\right) = \frac{E(-\Lambda^{-1}\left(\bar{y}_n\right)|\bar{x}_n) - \beta_0}{\beta_4} \tag{14}$$

Equation (14) implies that a linear transformation of $z\left(\bar{x}_n\right)$ can be identified from the data. Figure 8 shows the estimated $z\left(\bar{x}_n\right)$ for the Chicago CMSA. As the figure shows, $z\left(\bar{x}_n\right)$ must be substantially nonlinear if it is to generate the observed threshold in the regression relationship.

Because the distribution of $\bar{x}_n$ is observed for each city, $z\left(\bar{x}_n\right)$ also implies a distribution of z_i across individuals. Since z_i represents a worker's general productivity, the shape of that distribution should be plausible for productivity. Figure 9 shows the empirical distribution of college degrees across neighborhoods in the Chicago CMSA. As the figure shows, this distribution is positively skewed. Figure 10 shows the distribution of productivity across individuals implied by the estimated $z\left(\bar{x}_n\right)$ As the figure shows, the implied distribution of productivity is symmetric or negatively skewed.

[6] I show in Appendix A that this assumption is innocuous, but makes the algebra much clearer.

This implied distribution can be compared to independent estimates of the distribution of productivity across the economy. It is well known that the distribution of wages is not symmetric but is positively skewed. In a competitive labor market with no frictions, the distribution of wages and the distribution of marginal productivity are identical. However, a labor market without frictions is a strong assumption, especially in a study of unemployment. In the presence of costly search, workers are not necessarily paid their marginal product, so wages and productivity do not necessarily have identical distributions. Koning, Ridder, and van den Berg [13] estimate the distribution of productivity in the context of an equilibrium search model. They find a distribution which is highly positively skewed. This suggests that a productivity distribution should not be symmetric but rather skewed.

Could a positively skewed distribution of z_i generate a nonlinear $z(\bar{x}_n)$ like that shown in Figure 8? The sorting process described earlier implies that if a neighborhood's productivity level $\bar{z}_n$ is in percentile X of the productivity distribution, its average educational attainment $\bar{x}_n$ must be exactly in percentile X of its distribution:

$$F_x(\bar{x}_n) = F_z(\bar{z}_n) \tag{15}$$

Suppose both distributions are skewed as in Figures 9 and 10. When $\bar{x}_n$ is low, F_x' is high and F_z' is low. Relatively large percentile increases are associated with small increases in $\bar{x}_n$, while the opposite is true for $\bar{z}_n$. In the Chicago CMSA, the difference in $\bar{x}_n$ between the worst-educated neighborhood and the median neighborhood is about 15 percentage points, while the difference between the median and the best is about 85. In contrast, when $\bar{z}_n$ is in its lower two quartiles, F_z' is low - a relatively large change in $\bar{z}_n$ is needed to generate a moderate change in percentile terms. So a small increase in $\bar{x}_n$ leads to a larger increase in percentile terms which leads to an even larger increase in $\bar{z}_n$. When $\bar{x}_n$ is high, the reverse logic applies, generating a steep slope in $z(\bar{x}_n)$ when $\bar{x}_n$ is low and a flat slope when it is high. If both distributions are skewed in the same direction, there will be no threshold relationship. In particular, if the distributions have the same shape ($F_x(X) = F_z(aX)$ for some constant a), then the relationship implied by equation (15) will be exactly linear.

To summarize, a $z(\bar{x}_n)$ function which generates the observed relationship between unemployment and education without social interactions requires that unobserved productivity have a roughly sym-

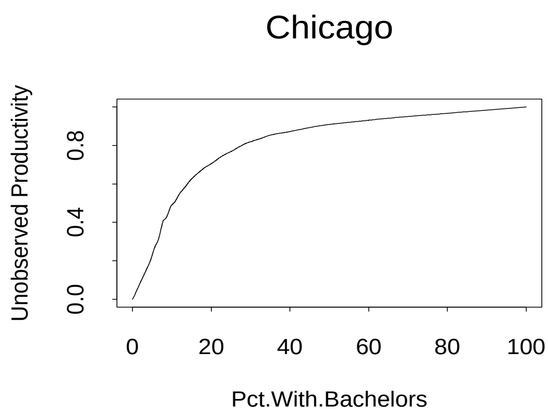

Figure 8: Estimated $z\left(\bar{x}_n\right)$ for Chicago. β_0 and β_4 set so that range of $z\left(\bar{x}_n\right)$ is $[0, 1]$.

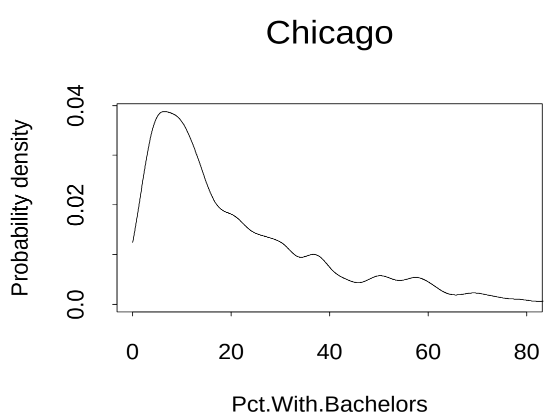

Figure 9: Distribution of $\bar{x}_n$ across neighborhoods in Chicago.

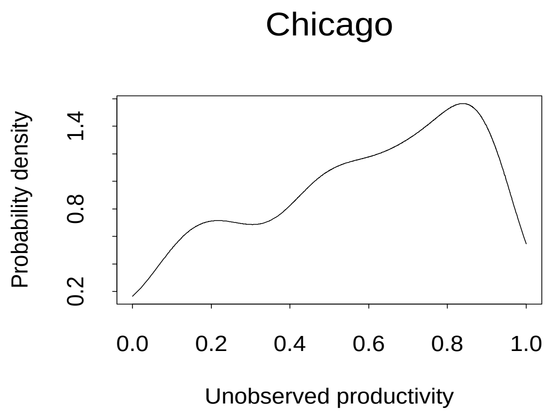

Figure 10: Distribution of z_i across individuals in Chicago, for estimated $z(\bar{x}_n)$.

metric or negatively skewed distribution, like that shown in Figure 10. A more plausibly shaped distribution for productivity is a positively skewed one similar to that shown in Figure 9, and this distribution generates a linear $z(\bar{x}_n)$, which is rejected by the data. I thus conclude that sorting alone does not generate the kind of thresholds observed in the data.

3.4 Contextual vs. endogenous effects

In this section I that assume there is no sorting effect, and evaluate whether contextual or endogenous neighborhood effects can plausibly generate the thresholds seen in the data.

To isolate the contextual effect, I assume that the endogenous and sorting effects are zero ($\beta_3 = \beta_4 = 0$). I also fix the direct effect of education at zero ($\beta_1 = 0$). Under these restrictions, equation (10) reduces to the following:

$$\bar{y}_n = \Lambda\left(-(\beta_0 + \beta_2 \bar{x}_n + \bar{\epsilon}_n)\right) \tag{16}$$

Equation (16) can be estimated from neighborhood data. Applying the log odds transformation to both sides produces:

$$\Lambda^{-1}(\bar{y}_n) = -(\beta_0 + \beta_2 \bar{x}_n + \bar{\epsilon}_n) \tag{17}$$

so a linear contextual effect translates into a linear relationship between neighborhood educational attainment and the log odds ratio transformation of the unemployment rate.

To isolate the endogenous effect, I assume that the contextual and sorting effects are zero ($\beta_2 = \beta_4 = 0$). These restrictions generate the following reduced form:

$$\begin{aligned} \bar{y}_n &= \bar{x}_n * \Lambda\left(-(\beta_0 + \beta_1 + \beta_3 \bar{y}_n + \bar{\epsilon}_n)\right) \\ &+ (1 - \bar{x}_n) * \Lambda\left(-(\beta_0 + \beta_3 \bar{y}_n + \bar{\epsilon}_n)\right) \end{aligned} \tag{18}$$

Note that equation (18) has $\bar{y}_n$ on both sides, so the implied relationship between $\bar{y}_n$ and $\bar{x}_n$ must be solved as a fixed point or equilibrium problem. Brock and Durlauf [2] analyze a class of models which includes equation (18) and prove that if β_3 is large enough, the resulting equilibrium correspondence $\bar{y}_n(\bar{x}_n)$ will exhibit multiple equilibria and threshold behavior in a critical range. While multiple equilibria preclude identification of the exact coefficients in equation (18), threshold behavior will also carry over to the log odds ratio transformation. Strong endogenous effects thus imply that the regression relationship $E(\Lambda^{-1}(\bar{y}_n)\,|\bar{x}_n)$ is nonlinear, while linear contextual effects imply it is linear. As a result, distinguishing between endogenous effects and linear contextual effects can thus be done by testing for nonlinearity in the log odds ratio.[7]

I apply the Härdle-Mammen test described earlier to test the null hypothesis that $E(\Lambda^{-1}(\bar{y}_n)\,|\bar{x}_n)$ is linear against the alternative that it is not.[8] I find that linearity is rejected for all cities. Combined with the results in the previous section, this provides evidence that an economically relevant endogenous effect is present.

4 Extensions

4.1 Race

Because this study is restricted to a single explanatory variable, missing variables are of substantial concern. In addition to the indirect methodology pursued in the previous section, it would be useful to control directly for variables that are known to affect employment

[7] This result was pointed out to me by William A. Brock.

[8] Note that this test is limited in that it can only reject a linear contextual effects model. It has no power to reject a nonlinear contextual effect or any type of endogenous effect.

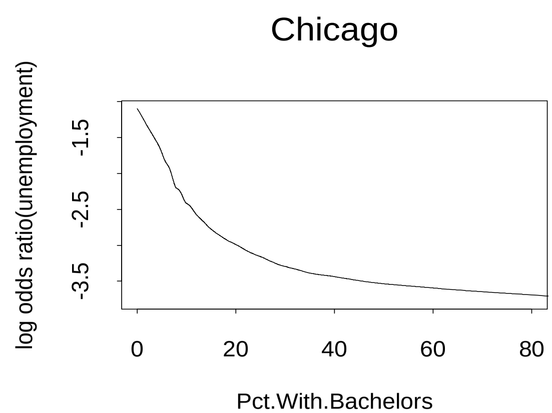

Figure 11: Estimated $E(\Lambda^{-1}(\bar{y}_n)|\bar{x}_n)$ for Chicago CMSA.

rates. Two candidate missing variables are age and race. Unemployment rates for the young and for African Americans are both substantially higher than average [11, Table 16.5]. If either of these variables is systematically related to the percentage of college graduates in a neighborhood, then this relationship could induce a spurious threshold.

Many of the standard methods of controlling for a given variable are problematic in this context. One common practice is to orthogonalize the dependent variable with respect to the other neighborhood characteristics and treat this residual as the quantity of interest. Under the hypothesis of endogenous social interactions, however, neighborhood unemployment enters into individual unemployment probabilities in a highly nonlinear way, creating a threshold. Because any other relevant variable would also have this nonlinear social multiplier effect, the standard practice is invalid precisely under the conditions of interest. In this section and in Section 4.2, I discuss two approaches which overcome the limitations of the standard practice.

To see how racial differences in employment patterns are sufficient to generate thresholds, consider the following example. There are no social interactions, and the only individual variables that matter for employment are race and education. The unemployment rates of the four groups are:

$$\Pr(y_i = 1|\text{black, college degree}) \quad = \quad 0.10$$

$$
\begin{aligned}
\Pr(y_i = 1|\text{black, no college degree}) &= 0.20 \\
\Pr(y_i = 1|\text{white, college degree}) &= 0.05 \\
\Pr(y_i = 1|\text{white, no college degree}) &= 0.15
\end{aligned}
$$

Suppose also that neighborhoods are completely racially segregated. Within neighborhoods of a particular race, the aggregate employment-education relationship is linear, as shown by the dotted lines in Figure 12. Now suppose that all neighborhoods with fewer than 10 percent college graduates are all black and all neighborhoods with more than 10 percent are all white. The resulting aggregate relationship between $\bar{x}_n$ and $\bar{y}_n$ will then look like the solid line in Figure 12, even though there are no social interaction effects. While this example is extreme, it illustrates that a relationship between percent black and average educational attainment in neighborhoods can lead to spurious inference of social interaction effects.

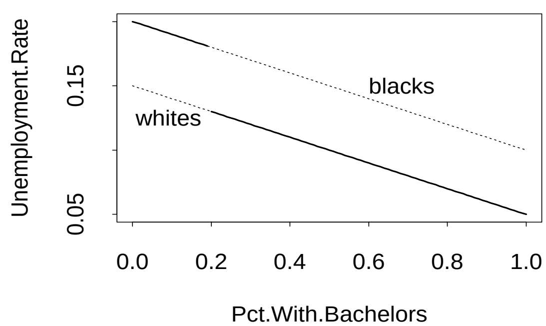

Figure 12: Example of spurious threshold due to ethnic group differences in unemployment rates.

Fortunately, the Census provides a breakdown of both education and employment status in a neighborhood by racial category. Let $\bar{x}_n^b$ be the fraction of black residents of neighborhood n who are college graduates, and $\bar{y}_n^b$ be their unemployment rate. Define $\bar{x}_n^w$ and $\bar{y}_n^w$ similarly for white residents. If the argument above is empirically relevant, then $E(\bar{y}_n^b|\bar{x}_n^b)$ and $E(\bar{y}_n^w|\bar{x}_n^w)$ will both be linear. Figure 13 shows the estimated regressions by racial category for the Chicago CMSA. Although unemployment rates are significantly higher among blacks, the threshold remains for both blacks and whites. The threshold remains for the fourteen other cities which had significant African-American populations. [9] In addition, I perform Härdle-Mammen linearity tests and find that linearity is rejected for all fourteen cities. Accounting for the impact of simple differences in neighborhood racial composition does not affect the results of Section 3.

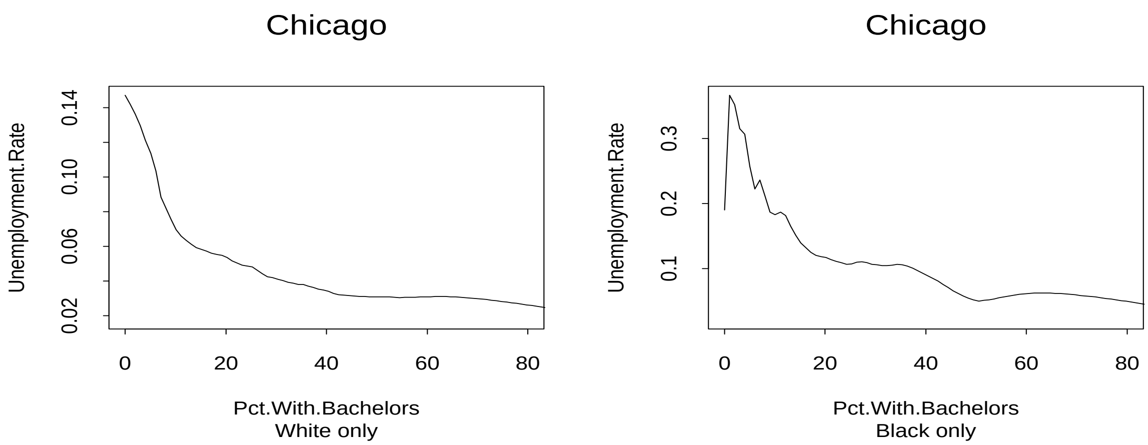

Figure 13: Nonparametric regressions by racial category. First graph shows estimated regression for whites, second shows regression for blacks.

4.2 Racial segregation

A related alternative hypothesis is that educational segregation is a proxy variable for racial segregation, and that racial segregation itself has negative effects on employment outcomes. The relative importance of segregation by race and by economic status in exacerbating black poverty is a matter of active current debate,[10] so this alternative form of social interaction effect should be considered. The distinction

[9] Six cities – Boston, Minneapolis - St. Paul, Phoenix, Tampa - St. Petersburg, San Diego, and Seattle – were dropped because they have few black residents.

[10] The work of Massey and Denton [16, and others] argues in favor of the view that racial segregation matters most, in contrast to Wilson's [21] focus on economic segregation.

between the employment effect of being black and the effect of living in a majority-black neighborhood is subtle, but can be investigated separately in this case. To do so, I split each city in my sample into those census tracts which are majority black and those which are not. The distribution of percent black in a neighborhood is bimodal in the sample, and almost all neighborhoods can be clearly categorized as having mostly black or mostly white residents. I estimate the basic regression $E(\bar{y}_n|\bar{x}_n)$ for each subsample separately. Again, six cities were dropped from the analysis. The results for the Chicago CMSA are shown in Figure 14. As the figure shows, the apparent threshold is much more prominent for majority-black neighborhoods. This lends some support to the argument that majority-black neighborhoods have unique characteristics. However, both thresholds remain, so the thresholds themselves are not an artifact of racial segregation. Again, Härdle-Mammen tests reject linearity for all fourteen cities.

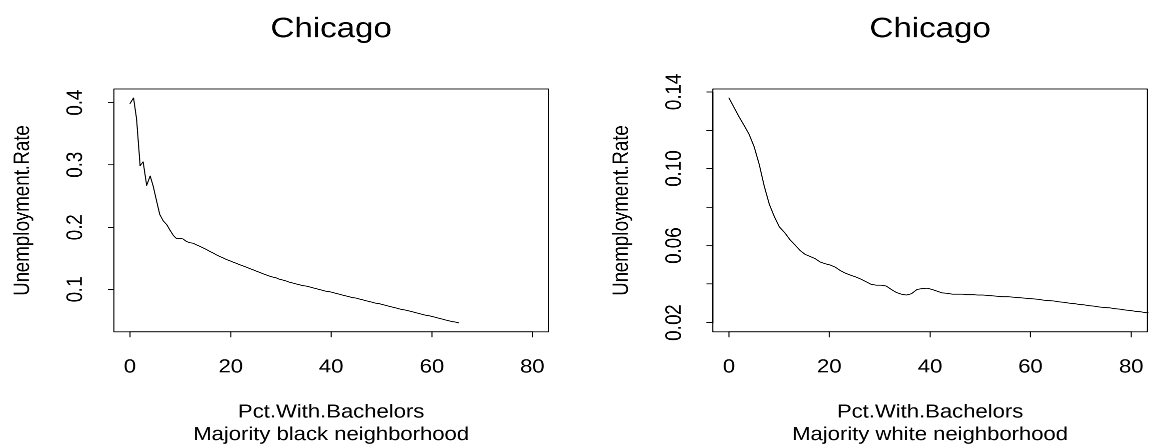

Figure 14: Nonparametric regressions by racial majority in neighborhood.

4.3 Age

The unemployment rate of individuals between 16 and 24 years of age is roughly twice that of those between 25 and 64. If the fraction of young adults in a community varies systematically with the educational attainment of its older members, the age distribution of the neighborhood could be an important missing variable. In addition, as education level is calculated for residents over age 25 and unemployment is calculated for residents over 16, these variables do not cover exactly the same populations as assumed in previous sections. Unemployment for those in the 25-64 age range would be a preferable dependent variable, but is unavailable at the tract level. To address this issue, I solve for bounds on the true regression function.

Let $\bar{y}_n^a$ be the unemployment rate for individuals in age range a and let p be the proportion of the labor force below age 25.

$$\bar{y}_n^{16-64} = \bar{y}_n^{25-64} * (1-p) + \bar{y}_n^{16-24} * p \tag{19}$$

Solving this equation for the variable of interest, the unemployment rate of 25-64 year-olds:

$$\bar{y}_n^{25-64} = \frac{\bar{y}_n^{16-64} - \bar{y}_n^{16-24} * p}{1-p} \tag{20}$$

The unemployment rate $\bar{y}_n^{16-64}$ and population proportion p are observed in the data for each tract, but the youth unemployment rate $\bar{y}_n^{16-24}$ is not. Because unemployment rates are higher for youth, the unemployment rate of 16-64 year-olds is an upwardly biased estimate of the unemployment rate of 25-64 year-olds.

While the data do not show unemployment rates for either age group, suppose there is a plausible upper bound on youth unemployment $\bar{y}_n^{16-24}$.

$$\bar{y}_n^{16-24} \leq \bar{y}_n^{\max}$$

As equation (20) is monotonic in $\bar{y}_n^{16-64}$, this restriction implies a lower bound on unemployment for adult workers.

$$\bar{y}_n^{25-64} \geq \frac{\bar{y}_n^{16-64} - \bar{y}_n^{\max} * p}{1-p} \tag{21}$$

The upper bound is simply the measured unemployment rate for the tract. Once the bounds are calculated for each tract, the desired nonparametric regression function $E(\bar{y}_n^{25-64}|\bar{x}_n)$ can be bounded:

$$E(\frac{\bar{y}_n^{16-64} - \bar{y}_n^{\max} * p}{1-p}|\bar{x}_n) \leq E(\bar{y}_n^{25-64}|\bar{x}_n) \leq E(\bar{y}_n^{16-64}|\bar{x}_n) \tag{22}$$

The constraint that the adult unemployment rate cannot be less than zero may bind as well. Provided that we set an upper bound on youth unemployment, the upper and lower bounds of equation (22) can be estimated.

Figure 15 shows estimated regression bounds for the Chicago CMSA. The first depicts the worst-case scenario of 100% youth unemployment, the second depicts bounds for 50% youth unemployment. Of the tracts in Chicago, only nine have unemployment rates higher than fifty percent, although the number with youth unemployment rates above fifty percent is likely to be higher. The appropriate interpretation of these bounds is different from that of a confidence interval. Any regression function that can be drawn between the upper and lower bound is consistent with the data, and there is no well-defined sense in which any such function is "more likely" than any other such function. As the figure shows, for the worst-case bounds, we cannot reject a linear relationship between neighborhood education and neighborhood unemployment. However, if the youth unemployment rate is no more than fifty percent in each tract, a linear relationship can be rejected.

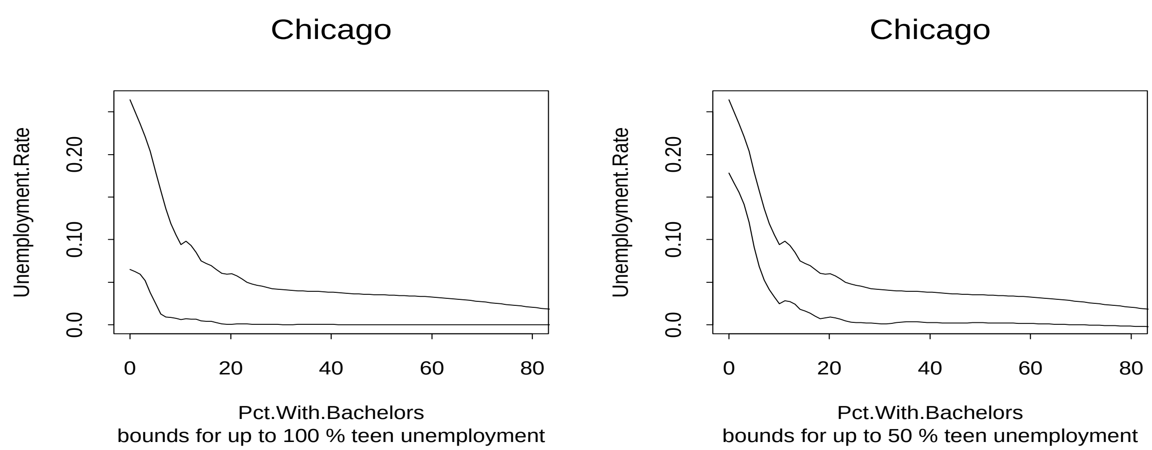

Figure 15: Bounds on regression function, Chicago CMSA. The first graph shows the worst-case bounds of 100% youth unemployment, the second shows an upper bound of 50% youth unemployment.

5 Conclusion

This paper provides evidence that there are thresholds in neighborhoods due to social interaction effects. Neighborhoods in which fewer than twenty percent of over-25 residents has a Bachelor's degree experience much larger unemployment rates. This stylized fact can be most easily explained with a model in which an endogenous neighborhood effect creates epidemic-style outcomes. While it can also be explained by a sorting process, doing so requires an implausible distribution of productivity.

The paper also shows that modelling the sorting process itself can be a valuable tool in solving the identification problems associated with empirical work on social interactions. Future work should produce a richer and more useful sorting model by exploiting information on the movements of families found in longitudinal data sets. Such a model could place further structure on the implications of requiring sorting alone to explain apparent social interaction effects.

References

[1] Roland Bénabou. Workings of a city: Location, education, and production. *Quarterly Journal of Economics*, 108:619–652, 1993.

[2] William A. Brock and Steven N. Durlauf. Discrete choice with social interactions I: Theory. Working Paper 5291, NBER, 1995.

[3] Jonathan Crane. The epidemic theory of ghettos and neighborhood effects on dropping out and teenage childbearing. *American Journal of Sociology*, 96:1226–1259, 1991.

[4] Steven N. Durlauf. A theory of persistent income inequality. *Journal of Economic Growth*, 1:75–93, 1996.

[5] William N. Evans, Wallace E. Oates, and Robert M. Schwab. Measuring peer group effects: A study of teenage behavior. *Journal of Political Economy*, 100:966–991, 1992.

[6] Lisa Finneran and Morgan Kelly. Labour market networks, underclasses and inequality. Working paper, University College Dublin, 1997.

[7] Edward L. Glaeser, Bruce Sacerdote, and José A. Scheinkman. Crime and social interactions. *Quarterly Journal of Economics*, 111(2):507–548, 1996.

[8] Mark Granovetter. *Getting a Job: A Study of Contacts and Careers.* University of Chicago Press, 1995. Second Edition.

[9] Wolfgang Härdle. *Applied Nonparametric Regression.* Cambridge University Press, 1990.

[10] Wolfgang Härdle and Enno Mammen. Comparing nonparametric versus parametric regression fits. *The Annals of Statistics*, 21(4):1926–1947, 1993.

[11] G.E. Johnson and P.R.G. Layard. The natural rate of unemployment: Explanation and policy. In O. Ashenfelter and R. Layard, editors, *Handbook of Labor Economics*, volume 2, chapter 16, pages 921–997. Elsevier Science, 1986.

[12] George L. Kelling and Catherine M. Coles. *Fixing Broken Windows: Restoring Order and Reducing Crime in Our Communities.* Simon and Schuster, 1996.

[13] Pierre Koning, Geert Ridder, and Gerard J. van den Berg. Structural and frictional unemployment in an equilibrium search model with heterogeneous agents. *Journal of Applied Econometrics*, 10:S133–S151, 1995.

[14] Brian V. Krauth. A dynamic model of job networks and persistent inequality. Working Paper 98-06-049, Santa Fe Institute, June 1998. Available at `http://www.ssc.wisc.edu/~bkrauth`.

[15] Charles F. Manski. Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60:531–542, 1993.

[16] Douglas S. Massey and Nancy A. Denton. *American Apartheid: Segregation and the Making of the Underclass.* Harvard University Press, 1993.

[17] Flynn McRoberts and Linnet Myers. Resettling the city's poor: Out of the Hole, into another. *Chicago Tribune*, August 23, 1998.

[18] James E. Rosenbaum and Susan Popkin. Employment and earnings of low-income blacks who move to middle-class suburbs. In Christopher Jencks and Paul Peterson, editors, *The Urban Underclass.* The Brookings Institution, Washington D.C., 1991.

[19] Thomas Schelling. *Micromotives and Macrobehavior.* W.W. Norton, 1978.

[20] Giorgio Topa. Social interactions, local spillovers, and unemployment. Working paper, New York University, 1997.

[21] William Julius Wilson. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy.* University of Chicago Press, 1987.

[22] William Julius Wilson. *When Work Disappears: The World of the New Urban Poor.* Alfred A. Knopf, 1996.

A The college wage premium

In Section 3, I assume that the parameter β_1, which describes the college wage premium, is equal to zero. This makes the estimation procedure much clearer. However, it is a somewhat arbitrary assumption. In this appendix I show that the results are not affected by the value of β_1.

With the assumption of no social interaction effects, equation (10 can be written as:

$$\bar{y}_n = \bar{x}_n * \Lambda\left(-(\beta_0 + \beta_1 + \beta_4 z\left(\bar{x}_n\right) + \bar{\epsilon}_n)\right) + (1 - \bar{x}_n) * \Lambda\left(-(\beta_0 + \beta_4 z\left(\bar{x}_n\right) + \bar{\epsilon}_n)\right) \tag{23}$$

Suppose we have a value for β_1. Define:

$$g_n(X) \equiv \bar{y}_n - \bar{x}_n * \Lambda\left(-(\beta_1 + X)\right) - (1 - \bar{x}_n) * \Lambda\left(-X\right) \tag{24}$$

Then for each neighborhood n, define Z as the solution to $g_n(Z) = 0$. The variable Z is a more general version of the log odds ratio transformation of $\bar{y}_n$ utilized in Section 3. Because g is a monotonic function which passes through zero, Z exists, is unique, and:

$$Z = \beta_0 + \beta_4 z\left(\bar{x}_n\right) + \bar{\epsilon}_n \tag{25}$$

Because $E(\bar{\epsilon}_n|\bar{x}_n) = 0$, the function $z\left(\bar{x}_n\right)$ can thus be estimated (up to a linear transformation) using:

$$z\left(\bar{x}_n\right) = E(Z|\bar{x}_n) \tag{26}$$

Figure 16 shows the estimated $z\left(\bar{x}_n\right)$ function for several values of β_1. As the figure shows, the estimates are not sensitive to alternative assumptions about β_1.

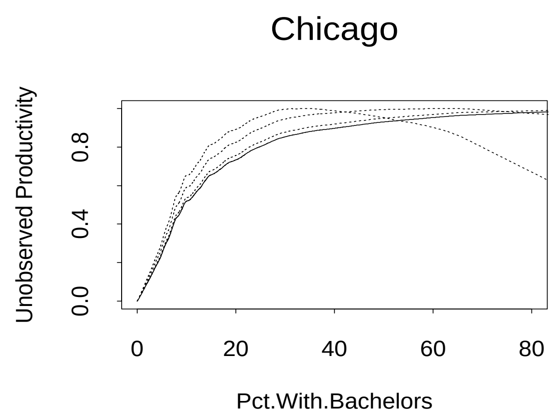

Figure 16: Estimated $z(\bar{x}_n)$ for Chicago CMSA, assuming values for β_1 of $\{0, 0.1, 0.5, 2\}$.