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

December 20, 2000

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 a wage distribution which is inconsistent with that observed in microeconomic data. Social interaction effects are thus a necessary element in any suitable explanation for the data.

JEL codes: C14, E24

Keywords: neighborhood effects, spillovers, networks, social interactions

*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, SUNY Albany, UCSD, Wisconsin, UBC, Toronto, the NBER Summer Institute, and the 8th World Congress of the Econometric Society.

1 Introduction

The distribution of unemployment across neighborhoods in most United States cities exhibits a great deal of spatial concentration. For example, Wilson [1996, p.19] finds that only one in three adults in Chicago's twelve poorest neighborhoods were working in a typical week in 1990, while roughly two in three were working in more affluent neighborhoods. Two possible explanations are typically advanced for this concentration. One explanation is that unemployment is a "social epidemic" in some neighborhoods. In other words, spillovers between neighbors in the effectiveness or cost of job search are large enough to produce either multiple equilibria (so that two otherwise identical neighborhoods may have different unemployment rates) or thresholds (so that two similar-but-not-identical neighborhoods may have very different unemployment rates). The other candidate explanation is that a person's neighborhood has minimal impact on employment and that instead individuals with poor employment prospects and individuals with strong employment prospects "sort" into different neighborhoods. This paper uses Census tract data from twenty large US cities to evaluate these two competing hypotheses.

Using Census tracts as a proxy for neighborhood, I find that average neighborhood unemployment (as predicted by a nonparametric regression) increases dramatically when the percentage of residents with college (Bachelor's) degrees falls below a critical value near twenty percent. In other words, there is a threshold relationship between neighborhood human capital and neighborhood unemployment, as predicted by the epidemic hypothesis. However, sorting of workers into neighborhoods may also explain this stylized fact. To distinguish empirically between the effects of sorting and of spillovers, I develop and estimate a structural model which includes residential sorting but does not include spillovers. The results indicate that, while sorting can in principle explain just about any relationship between neighborhood variables, explaining the particular empirical relationship between neighborhood human captial and unemployment also implies a negatively skewed distribution of wages and productivity across workers. This implication contrasts with the positively skewed distribution found in the microeconomic literature. The sorting model is thus inconsistent with the facts unless we allow for quantitatively important social interaction effects.

The idea of social epidemics, formalized in economics by Brock and Durlauf [2000] and described in Section 2.1 of this paper, appears repeatedly in debates on public policy. For example, criminologist

George Kelling and political scientist James Q. Wilson have seen their "broken windows theory" of crime, which holds that local crime rates respond sharply to small increases in disorder, implemented by New York City's police department in the 1990's. The policy changes (cracking down on small "quality of life" offenses) have received much attention in the press and given a great deal of credit for the rapid decline in crime rates that followed [1996]. Sociologist William Julius Wilson's well-cited analysis of unemployment and social pathology in Chicago's ghettos [1987,1996] argues that the problems in inner cities are best explained as an epidemic. A recent popularization of the idea of social epidemics by journalist Malcolm Gladwell [2000] has even spent time on the best-seller lists. However, systematic evidence on the empirical relevance of social epidemics is currently limited. Crane [1991] finds some evidence for epidemics in out-of-wedlock childbearing and high school dropout, but does not deal with sorting. Topa [2000] analyzes spillovers in unemployment between but not within Census tracts. Glaeser, Sacerdote, and Scheinkman [1996] use the variance in aggregate crime rates to estimate the strength of spillovers, accounting for the effects of sorting. However, their model rules out multiple equilibria by construction. This paper, in contrast, combines a focus on social epidemics with an analysis of sorting.

2 Do neighborhoods have thresholds?

As a first step, I estimate a nonparametric regression $E(\bar{u}_n|\bar{c}_n)$ of a neighborhood's unemployment rate ($\bar{u}_n$) on its average human capital ($\bar{c}_n$). 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 A model with thresholds

First I describe a stylized model of social interactions to show how threshold effects or multiple equilibria can arise. This model is closely related to that in Brock and Durlauf [1995], though the application to unemployment is specific to this paper. Individuals are exogenously assigned to neighborhoods, and are characterized by exogenous education level c_i. Each worker receives a wage offer (net of some reservation wage) w_i:

$$w_i = \beta_{int} + \beta_c c_i - \beta_u \bar{u}_n + \epsilon_i \quad (1)$$

where ϵ_i is independently and identically distributed and drawn from the distribution $F_\epsilon(.)$. The coefficient on $\bar{u}_n$ corresponds to the neighborhood effect. This neighborhood effect could be the result of some conformity effect, informational externalities, or any other externality. 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.

$$u_i = \begin{cases} 0 & \text{if } w_i \geq 0 \\ 1 & \text{if } w_i < 0 \end{cases} \tag{2}$$

This yields the following model:

$$E(u_i|c_i, \bar{u}_n) = F_\epsilon\left(-(\beta_{int} + \beta_c c_i - \beta_u \bar{u}_n)\right) \tag{3}$$

In a large neighborhood, we have:

$$\bar{u}_n = \bar{c}_n F_\epsilon\left(-\beta_{int} - \beta_c + \beta_u \bar{u}_n\right) + (1 - \bar{c}_n) F_\epsilon\left(-\beta_{int} + \beta_u \bar{u}_n\right) \tag{4}$$

Notice that $\bar{u}_n$ appears on both sides of this equation. Assume a normal or logistic distribution for the residual and define an equilibrium as a level of neighborhood unemployment which is consistent with equation (4). For sufficiently large values of β_u, Brock and Durlauf show there exist multiple equilibria over some range of $\bar{c}_n$. In addition, the equilibrium correspondence will be discontinuous. An numerical example of the equilibrium relationship between $\bar{u}_n$ and $\bar{c}_n$ for a particular set of parameter values is shown in Figure 1. As the figure shows, models like this can produce thresholds in $E(\bar{u}_n|\bar{c}_n)$. This example, while highly stylized and not necessarily appropriate for empirical application, shows that simple models with social interaction effects can generate thresholds.

2.2 Data

The data source for this study is the 1990 United States Census (Summary Tape File 3A). Each data point is a census tract, a geographic region with population usually between 2,500 and 8,000 residents which provides a rough approximation to neighborhood. Topa [2000] also studies the distribution of unemployment across urban census tracts, though his focus is on the patterns of correlation in unemployment between neighboring tracts rather than patterns within tracts. The census tracts in the sample are drawn from the five largest cities in each of the four major regions of the United States. Neighborhood unemployment rates, which will be denoted by $\bar{u}_n$, are calculated for civilians in the labor force. Neighborhood human capital, denoted by $\bar{c}_n$, is measured as the fraction of neighborhood residents over age 25 who have Bachelor's degrees. A Census tract is omitted if it has fewer than 300 residents in the labor force.

One might wish to know how well the Census tract approximates a neighborhood. The tracts are initially drawn by the Census Bureau to be relatively homogeneous with respect to economic and demographic characteristics, and to reflect known neighborhood configurations. While this makes them a reasonable proxy for neighborhood, there is a potential for endogeneity problems. When deciding whether to put a group of families into one of two adjoining tracts with different levels of economic success, this procedure would tend to put a poor family into the "poor" tract and a wealthy family into the "wealthy" tract. However, the boundaries of many urban tracts were drawn several decades before 1990, and are infrequently adjusted, so the endogeneity of the boundary is unlikely to be a significant issue. Another potential problem is that 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. If this is the case, then any use of census tracts to represent neighborhoods will tend to understate the importance of social interactions.

The cities in the sample represent the five largest Metropolitan Statistical Areas (MSA) or Consolidated Metropolitan 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). For simplicity I will refer to a given MSA or CMSA simply as a "city". In the analysis, each city will be treated as a single labor market, with its own joint distribution of wages, education, and other characteristics. Table 1 shows the mean and median values of the two variables, as well as the number of observations in each city. As the table shows, there is a substantial variation in both average education levels and unemployment rates across cities.

The results reported in this paper are robust to a number of alternative decisions about the data. In particular using male unemployment or female unemployment in place of overall unemployment, using the employment/population ratio instead of the unemployment rate, and treating each county or PMSA as a separate labor market all produce similar results to those reported here.

2.3 Identification

The empirical importance of social interaction effects is controversial among social scientists. Manski [2000] argues that this disagreement is the inevitable result of a fundamental identification problem. 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. As Manski points out, there is no generally applicable way to determine the relative importance of these two factors from nonexperimental data on behavior as long as the reference group is selected by purposeful economic agents rather than by random experiment.

Unfortunately, social interaction effects and sorting have very different policy implications. For example, if there are true neighborhood effects in unemployment, the residential distribution of individuals has an impact on individual employment outcomes. Relatively expensive policies such as the Chicago Housing Authority's attempts to relocate public housing tenants into more economically diverse neighborhoods [McRoberts 1998] 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. These programs will be ultimately ineffective without 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 [Granovetter 1995]. Other studies indicate that employment patterns of one's social environment have significant impacts on attitudes towards work. For example, the field work of Wilson [1996] finds that young people in high-unemployment neighborhoods have much greater pessimism about the labor market and much weaker understanding of what prospective employers want. Both the networking effect and the attitudinal effect imply that living in a high-unemployment social environment has a detrimental effect on one's own probability of employment.

However, economists are historically suspicious of attitudinal surveys. Furthermore, 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:

$$\Pr(y_i = 1|x_i, \bar{z}_n) = F_\epsilon\left(\beta x_i + \gamma \bar{z}_n\right) \tag{5}$$

where y_i is the outcome, x_i is a vector of individual characteristics, $\bar{z}_n$ is some variable describing neighborhood composition, and $F_\epsilon(.)$ is either the normal or logistic CDF. It is tempting to say that the coefficient γ measures the social interaction effect, but it actually measures the combined social interaction effect and sorting effect. Because of the identification problem, additional assumptions are needed to estimate the two effects separately.

The early literature assumes exogenous selection into neighborhoods. Under exogenous selection, there can be some degree of sorting on the regressors, but there can be no sorting on the outcome or any component of the unobserved term. In this case, there is no sorting effect (though there may be sorting), and γ actually measures the social interaction effect. For most economically interesting questions, exogenous selection is highly unlikely. It would be surprising if dropping out of school, commiting crimes, receiving public aid, entering or exiting employment, or out-of-wedlock childbearing did not affect neighborhood choice. If the outcome itself affects neighborhood choice, an econometrician estimating equation (5) must address this endogeneity problem.

The more recent literature on social interaction effects [Evans et al. 1992, Glaeser et al. 1996, Rosenbaum 1991] uses instrumental variables and natural experiments to dispense with the unpalatable assumption of exogenous selection. However, usable instruments or natural experiments are rare, and those which have been used are themselves controversial. This paper follows a different approach by developing a parametric model of the sorting process. Under the assumption of no social interactions, one can estimate certain parameters and compare them with what is known from the microeconomic literature. If fitting the model to the data contradicts known results from other studies, we can must reject one of the assumptions built into the model, including possibly the assumption of no social interactions. The details and results of this approach are outlined in Section 3.

2.4 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 all neighborhoods in a city, not just a sample. Both of these features are exploited in this study to achieve very high precision in estimates and to model the neighborhood formation process itself. These benefits of aggregate data come at the cost of inability to identify individual-level parameters.

The standard method for ascertaining the existence of social interaction effects under the assumption of exogenous selection is to estimate equation (5) using individual-level data and test the null hypothesis that the coefficient γ is zero. In a nonparametric context, the analogous procedure would be to estimate:

$$E(u_i|c_i, \bar{c}_n) \tag{6}$$

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

$$H_0 : E(u_i|c_i, \bar{c}_n = X) = E(u_i|c_i, \bar{c}_n = X') \quad \text{for all } \{X, X'\} \tag{7}$$

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

$$\begin{aligned} E(\bar{u}_n|\bar{c}_n = X) &= (1 - \bar{c}_n) E(u_i|c_i = 0, \bar{c}_n = X) \\ &+ \bar{c}_n E(u_i|c_i = 1, \bar{c}_n = X) \end{aligned} \tag{8}$$

Under the null hypothesis (7), Equation (8) is linear in $\bar{c}_n$. In other words, as long as c_i is a single binary variable, testing for the linearity of $E(\bar{u}_n|\bar{c}_n)$ is analogous to using individual data to test $\gamma = 0$ in equation (5) under the assumption of exogenous selection, the primary exercise in much of the social interactions literature.[1]

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 2) from one that has a steep slope over a short range of the explanatory variable (Figure 3). As a result,

[1] If $E(u_i|c_i, \bar{c}_n)$ is linear in $\bar{c}_n$, then $E(\bar{u}_n|\bar{c}_n)$ will also be linear. The aggregate test will generally fail to reject the null hypothesis (7), even though it is false and would be rejected by the individual-level test. This lack of power is unimportant to the results here, however, as the null hypothesis of linearity is rejected quite easily.

any continuous regression function that shows a large change in unemployment over a small range of neighborhood average human capital can be said to 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.

2.5 Results

Figure 4 shows a scatter plot of neighborhood unemployment ($\bar{u}_n$) versus neighborhood educational attainment ($\bar{c}_n$) 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 5 through 8 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 (see Härdle [1990], page 181), and the 95 percent confidence intervals shown are estimated by the bootstrap with 1,000 iterations.[2] Alterntive smoothing procedures produce similar results. As appears in 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 the social epidemic hypothesis.

As shown in Section 2.4, a statistical test for linearity in these regressions will be equivalent to a statistical test for the joint null hypothesis of exogenous selection and no social interactions. I implement this test using the nonparametric method proposed by Härdle and Mammen [1993]. 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

[2] To describe the procedure in more detail, I estimate an $X\%$ pointwise interval, where X is high enough that 95% of the bootstrapped regression functions are entirely within these intervals (see Härdle [1990] for a discussion of this procedure). As a result, the confidence interval has a 95% probability of containing the entire true regression function.

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 $E(\bar{u}_n|\bar{c}_n)$ is easily rejected by the data in all cases. 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; in Section 3 we attempt to distinguish between these two explanations.

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 5 - 8 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 thus a robust characteristic of the reduced form relationship between neighborhood resources and neighborhood unemployment. While this reduced-form result is consistent with epidemic models of social interaction effects, other structural models may have similar reduced-form implications. Section 3 considers alternative explanations and finds that the epidemic explanation is more plausible.

2.6 Alternative explanations: Race

The indirect methodology in Section 3 is aimed at determining in general whether missing variables could be the source of the apparent threshold. It is also possible to investigate more directly and under weaker assumptions whether some specific missing variables could be the source. In particular, unemployment rates for the young and for African Americans are both substantially higher than average (see for example Johnson and Layard [1986], 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.

Consider the following example in which 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:

$$\begin{aligned}
\Pr(u_i = 1|\text{black, college degree}) &= 0.10 \\
\Pr(u_i = 1|\text{black, no college degree}) &= 0.20 \\
\Pr(u_i = 1|\text{white, college degree}) &= 0.05 \\
\Pr(u_i = 1|\text{white, no college degree}) &= 0.15
\end{aligned}$$

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 9. Now suppose that neighborhoods are racially segregated, all neighborhoods with fewer than 20 percent college graduates are all black and all neighborhoods with more than 20 percent are all white. The resulting aggregate relationship between $\bar{c}_n$ and $\bar{u}_n$ will then look like the dark line in Figure 9, even though there are no social interaction effects. While this example is extreme, it illustrates that a relationship between racial composition and average educational attainment in neighborhoods can lead to spurious inference of social interaction effects.

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 the resulting 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.

Fortunately, the Census provides a breakdown of both education and employment status in a neighborhood by racial category allowing us to directly test whether differences in neighborhood racial composition explain the unemployment threshold. Let $\bar{c}_n^b$ be the fraction of black residents of neighborhood n who are college graduates, and $\bar{u}_n^b$ be their unemployment rate. Define $\bar{c}_n^w$ and $\bar{u}_n^w$ similarly for white residents. If the argument above is empirically relevant, then $E(\bar{u}_n^b|\bar{c}_n^b)$ and $E(\bar{u}_n^w|\bar{c}_n^w)$ will both be linear. Figure 10 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. [3] 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.

2.7 Alternative explanations: Age

A similar approach can indicate the importance of variations in a neighborhood's age distribution. 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, and could explain the unemployment threshold. In addition, as education level is calculated for residents over age 25 and unemployment is calculated for residents over 16, these variables do not describe the same population. 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 find nonparametric bounds [Manski 1994] on the true regression function.

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

$$\bar{u}_n = (1-q)\bar{u}_n^{25-64} + q\bar{u}_n^{16-24} \tag{9}$$

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

$$\bar{u}_n^{25-64} = \frac{\bar{u}_n - q\bar{u}_n^{16-24}}{1-q} \tag{10}$$

The unemployment rate $\bar{u}_n$ and population proportion q are observed in the data for each tract, but the youth unemployment rate $\bar{u}_n^{16-24}$ 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 separate unemployment rates for either age group, suppose there is a plausible upper bound $u^{\max}$ on youth unemployment. The "worst case" upper bound is 100 percent.

$$\bar{u}_n^{16-24} \le u^{\max}$$

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

As equation (10) is monotonic in $\bar{u}_n$, this restriction implies a lower bound on unemployment for adult workers.

$$\bar{u}_n^{25-64} \geq \frac{\bar{u}_n - (qu^{\max})}{1-q} \tag{11}$$

The constraint that the adult unemployment rate cannot be less than zero provides an additional lower bound. As we can reasonably expect that the under-25 workers in a particular neighborhood are at least as likely to be unemployed as the older workers, the upper bound on over-25 unemployment is simply the measured unemployment rate for the tract. Once the bounds are calculated for each tract, the desired nonparametric regression function $E(\bar{u}_n^{25-64}|\bar{c}_n)$ can be bounded.

$$E(\bar{u}_n^{25-64}|\bar{c}_n) \in \left[E\left(\max\left\{0, \frac{\bar{u}_n - (qu^{\max})}{1-q}\right\} \middle| \bar{c}_n \right), E(\bar{u}_n|\bar{c}_n) \right] \tag{12}$$

Once a value for $u^{\max}$ is set, the upper and lower bounds of equation (12) can be estimated.

Figure 11 shows estimated regression bounds for the Chicago CMSA. The first graph depicts the “worst case” bounds for the scenario of 100% youth unemployment, the second depicts “reasonable restriction” bounds for 50% youth unemployment. The appropriate interpretation of the 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. 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. Because there are hundreds of Census tracts in Chicago, the assumption of an upper bound of 50% on youth unemployment is quite reasonable. It is thus unlikely that a bias due to age composition will explain the threshold.

3 Addressing the sorting hypothesis

The results in the previous section established that either sorting effects or social interaction effects must be present. This section applies a new strategy to distinguish between the two types of effects. The basis of this strategy is to recognize that the critical question is not the presence or absence of sorting effects. Indeed, basic economics implies that there must be some sorting effect present for this problem. Instead, the question is whether or not we need to infer an economically significant degree of social interaction effects in order to explain the data. Section 3.2 develops a model of sorting. Section 3.3 shows that this model cannot explain the presence of a threshold in the data without having implications that are difficult to reconcile with reality. Because the model with social interactions described in Section 2.1 can explain the presence of a threshold without any particular assumption on the nature of the sorting effect, I conclude that, while sorting effects are also likely, social interaction effects are necessary to explain the threshold.

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 educational attainment is an explanatory variable. Another alternative, that individuals sort on employment status itself, can be easily ruled out. If individuals sort directly on employment status, the distribution of neighborhood unemployment rates should have many neighborhoods with no unemployment and a few neighborhoods with unemployment of 100 percent. As nearly all of the neighborhoods in the sample have unemployment rates which are strictly between zero and thirty percent, direct sorting on employment status is not a reasonable modeling assumption.

Some measure of income is the most promising candidate sorting variable. In a simple housing market with no externalities, neighborhood stratification on income level will occur if the most attractive 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.[4] Epple and Sieg [1999] find that a

[4] In the presence of quantitatively important externalities such as social interaction effects, the neighborhood formation problem becomes more complex because families may care who their neighbors are. Bénabou [1993] and Durlauf [1996] describe conditions under which social interaction effects reinforce the incentive to sort.

simple income-sorting model modified for some taste heterogeneity is consistent with the distribution of income within and across neighborhoods in the Boston area. The Census data also indicate that response to transitory income shocks is small. 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 imply that families do not always change residential location in response to a transitory change in income or employment.

In the remainder of Section 3, I use an indirect approach to exploit the idea that the primary 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 portable and somewhat permanent.

3.2 A model of sorting

This section develops the model with sorting. In the model each individual's employment outcome is described by a simple binary choice model and individuals sort into neighborhoods on an unobserved productivity variable. Each worker in a given city goes through the following process:

1. The worker draws an exogenous educational attainment $c_i \in \{0,1\}$ and productivity $p_i \in [0,1]$ from the joint distribution function $F_{p,c}(.)$. The conditional distribution $F_{p|c}(.)$ has a strictly monotone likelihood ratio, i.e., an individual with a college degree tends to have higher productivity.

2. The worker chooses a neighborhood based on a simple sorting rule. Workers stratify perfectly into N neighborhoods of size s based on their value of p_i. Ties are broken by lottery.

3. The worker receives a random wage offer. The net wage offer (wage offer minus reservation wage) is linear in education and productivity, and subject to both neighborhood level and idiosyncratic shocks ($\bar{\theta}_n$ and ϵ_i respectively):

$$w_i = \beta_{int} + \beta_c c_i + \beta_p p_i + \bar{\theta}_n + \epsilon_i$$

4. The worker accepts the net wage offer if it is positive:

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

The variables $\bar{\theta}_n$ and ϵ_i represent unobserved neighborhood and individual level shocks, respectively. Unlike the unobserved productivity variable p_i, which is known by the worker before choosing location, the shocks $\bar{\theta}_n$ and ϵ_i affect the worker after the locational choice has been made. Let $\bar{c}_n$ represent the neighborhood average of c_i, and define $\bar{u}_n$ and $\bar{p}_n$ similarly. Assume the neighborhood-level shock has a mean of zero conditional on the other neighborhood characteristics:

$$E(\bar{\theta}_n|\bar{c}_n) = 0 \tag{13}$$

Also assume that ϵ_i is independent of individual and neighborhood characteristics, and independently and identically distributed across individuals with distribution function $F_\epsilon\left(.\right)$.

Under these assumptions, the unemployment probability of an individual follows a standard discrete choice model:

$$\Pr(u_i = 1|c_i, p_i, \bar{c}_n, \bar{\theta}_n) = F_\epsilon\left(-(\beta_{int} + \beta_c c_i + \beta_p p_i + \bar{\theta}_n)\right) \tag{14}$$

where $F_\epsilon\left(X\right) \equiv \Pr(\epsilon_i \leq X)$. Equation (14) includes the individual-level variables c_i and p_i, while the data consist of neighborhood-level educational attainment and unemployment rates. In order to perform estimation, it is necessary to rewrite equation (14) in terms of aggregates.

Because of the MLRP assumption, $E(p_i|\bar{c}_n)$ is a strictly monotonic function. Let $p\left(\bar{c}_n\right) = E(p_i|\bar{c}_n)$. As the number of neighborhoods and the size of a typical neighborhood both go to infinity, neighborhoods become homogeneous in productivity. Every worker in a neighborhood with a fraction $\bar{c}_n$ of college graduates has productivity level $p_i = p\left(\bar{c}_n\right)$. Equation (14) becomes:

$$\Pr(u_i = 1|c_i, \bar{c}_n, \bar{\theta}_n) = F_\epsilon\left(-(\beta_{int} + \beta_c c_i + \beta_p p\left(\bar{c}_n\right) + \bar{\theta}_n)\right) \tag{15}$$

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

As the size of neighborhoods approaches infinity, $\bar{u}_n$ converges to $E(\bar{u}_n|\bar{c}_n, \bar{\theta}_n)$. Although the data naturally have finite neighborhoods, we can use an asymptotic approximation because the standard deviation of an average of a few thousand independent Bernoulli random variables is very 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. Using the approximation that $\bar{u}_n = E(\bar{u}_n|\bar{c}_n, \bar{\theta}_n)$ we can derive the following aggregate model:

$$\begin{aligned} \bar{u}_n &= \bar{c}_n * F_\epsilon\left(-(\beta_{int} + \beta_c + \beta_p p\left(\bar{c}_n\right) + \bar{\theta}_n)\right) \\ &+ (1 - \bar{c}_n) * F_\epsilon\left(-(\beta_{int} + \beta_p p\left(\bar{c}_n\right) + \bar{\theta}_n)\right) \end{aligned} \tag{16}$$

In order to simplify the algebra, assume the direct benefit of education (β_c) is zero. The results are not sensitive to alternate values of β_c.[5] Equation (16) reduces to the following:

$$\bar{u}_n = F_\epsilon\left(-(\beta_{int} + \beta_p p\left(\bar{c}_n\right) + \bar{\theta}_n)\right) \tag{17}$$

Taking $E(F_\epsilon^{-1}\left(.\right)|\bar{c}_n)$ of both sides of equation (17) and rearranging we get:

$$p\left(\bar{c}_n\right) = \frac{E(-F_\epsilon^{-1}\left(\bar{u}_n\right)|\bar{c}_n) - \beta_{int}}{\beta_p} \tag{18}$$

Equation (18) implies that a linear transformation of $p\left(\bar{c}_n\right)$ can be identified from the data. Let $F_p\left(.\right)$ represent the distribution of p_i accross individuals, and let $F_{\bar{c}}\left(.\right)$ represent the distribution of neighborhood average education across neighborhoods. The sorting process implies that:

$$F_p\left(p\left(\bar{c}_n\right)\right) = F_{\bar{c}}\left(\bar{c}_n\right) \tag{19}$$

or:

$$F_p\left(\frac{E(-F_\epsilon^{-1}\left(\bar{u}_n\right)|\bar{c}_n) - \beta_{int}}{\beta_p}\right) = F_{\bar{c}}\left(\bar{c}_n\right) \tag{20}$$

This equation allows us to estimate the distribution of productivity ($F_p\left(.\right)$) using the joint distribution of $\bar{u}_n$ and $\bar{c}_n$ in the Census data, as well as some assumption about the distribution of ϵ_i. In the empirical work I assume that ϵ_i has a logistic distribution, so that $F_\epsilon\left(X\right) = \frac{e^X}{1+e^X}$. Section 3.5 evaluates the robustness of the results to assuming other distributions.

If we take the model literally, the productivity variable has several interpretations - permanent income, predicted wage (net of the reservation wage, college wage premium, and any idiosyncratic shocks), neighborhood average wage, etc. Although it is called a "productivity" variable, there is no requirement here that p_i is equal to a worker's marginal productivity, only that it is equal to

[5] This is shown in the notes for referees.

the worker's typical wage. This model is thus consistent with both efficient and inefficient labor markets, and also labor markets with features such as unions, racial discrimination, nepotism, and affirmative action. If we assume that reservation wages are constant across individuals (or linear in the expected wage), then the distribution of productivity implied by the sorting model should correspond to the distribution of wage offers conditional on education. Although the positive skewness of wage distributions is a well-known stylized fact in labor economics, estimates of the distribution of wage offers are uncommon. Koning, Ridder, and van den Berg [1995] estimate the distribution of of a productivity variable very similar to the variable p_i in this model using an equilibrium search model. They too find a distribution which is highly positively skewed. This suggests that we should expect the estimated distribution of productivity from the sorting model to have a strong positive skew as well.

Once we have an estimate of the distribution of productivity across individuals, this distribution can then be compared to independent estimates of the distribution of "productivity" across individuals. If the estimated $F_p(.)$ differs in some important way from these independent estimates, then the sorting model described here should be rejected as inconsistent with the data.

3.3 Results

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, and the results in other cities (available from the author) are quite similar. Figure 12 shows the estimated $p(\bar{c}_n)$ for the Chicago CMSA. As the figure shows, $p(\bar{c}_n)$ must be substantially nonlinear if it is to generate the observed threshold in the regression relationship. Figure 14 shows the distribution of productivity across individuals implied by the estimated $p(\bar{c}_n)$. As the figure shows, the implied distribution of productivity is negatively skewed or perhaps symmetric. As argued in the previous section, micro studies imply the opposite; the distribution of p_i should be positively skewed.

In order to see how this negative skewness is required to explain the data, it is necessary to explore how sorting could generate a threshold. Figure 13 shows the empirical distribution of average education across neighborhoods in the Chicago CMSA. As the figure shows, this distribution is positively skewed. Could a positively skewed distribution of p_i generate a nonlinear $p(\bar{c}_n)$ like that shown in Figure 12?

Equation (19) implies that if a neighborhood's productivity level $\bar{p}_n$ is in percentile X of the productivity distribution, its average educational attainment $\bar{c}_n$ must be exactly in percentile X of its distribution. Taking derivatives with respect to $\bar{c}_n$ we can rewrite this as:

$$p'(\bar{c}_n) = \frac{f_{\bar{c}}}{f_p} \tag{21}$$

where $f_{\bar{c}}$ is the probability density function (PDF) for the distribution of $\bar{c}_n$ and f_p is the PDF for the distribution of productivity. If both distributions are identical in shape, then $\frac{f_{\bar{c}}}{f_p}$ will be constant; equivalently, $p(\bar{c}_n)$ will be linear. If the distributions have quite different shapes, then $p'(\bar{c}_n)$ will vary, and $p(\bar{c}_n)$ will be nonlinear. In order to get the particular form of nonlinearity found in Figure 12, it is necessary for $\frac{f_{\bar{c}}}{f_p}$ to be high when $\bar{c}_n$ is low and low when $\bar{c}_n$ is high. In other words, one would need for $\bar{c}_n$ to be positively skewed and p_i to be negatively skewed.

To put the argument on a more intuitive footing, how could one explain an apparent threshold in $E(\bar{u}_n|\bar{c}_n)$ at $\bar{c}_n = 0.2$ using sorting alone? Suppose that a large fraction of neighborhoods have $\bar{c}_n$ around 0.2, with a smaller fraction spread around the other possible values. In that case a neighborhood with nineteen percent college graduates would nearly be one of the worst (in terms of whatever characteristics families sort on) and a neighborhood with twenty-one percent college graduates would nearly be one of the best. One would be quite surprised not to find a large difference in unemployment rates between these two neighborhoods, even in the absence of social interaction effects. [6] The only way this would not happen is if the sorting variable had a similar distribution across neighborhoods, so that the near-best and near-worst neighborhoods were fairly similar.

Sorting would produce a threshold effect in similar manner here. Because of its positive skew, the distribution of $\bar{c}_n$ across neighborhoods would tend to produce a threshold, unless the distribution of p_i were also positively skewed. But because micro studies imply that the distribution of p_i should be positively skewed, sorting alone will be unable to explain the data in a manner consistent with these other studies.

[6] Incidentally, this argument calls into serious question the results of Crane. He finds that a "percent high status workers in neighborhood" variable has a threshold effect (around 5% high-status workers) on the probability of school dropout and out-of-wedlock childbirth. However, this variable is very highly skewed, with the majority of neighborhoods having less than 10% high-status workers. As a result, it would be surprising if Crane did not find a threshold.

3.4 Implications

Rejection of the sorting model does not imply the absence of sorting. Common sense and economic theory would both indicate that income and employment have an important effect on choice of residence. Instead, rejection of the sorting model implies that some assumption of the model was incorrect. The sorting model entails several potentially strong assumptions, so any one of them may in principle prevent the model from fitting the data. There are three classes of assumptions to consider - functional form restrictions, assumptions about the sorting process itself, and the absence of a neighborhood effect.

The main functional form assumptions that may be questioned are the linearity of the wage function and the logistic distribution assumed for the individual residual ϵ_i. The linearity assumption is much weaker than first appears. Assume instead that individuals sort on productivity and that productivity enters into wages in a manner which is additively separable from education, but potentially nonlinear.

$$w_i = f_1(c_i) + f_2(p_i) + \bar{\theta}_n + \epsilon_i \tag{22}$$

As long as f_2 is monotonic, we can redefine $p_i = f_2(p_i)$. Equation (22) is linear in both education and the new productivity variable, p_i is a random variable that has a joint distribution with c_i that exhibits the MLRP property, and neighborhoods are perfectly sorted on p_i, so both the estimation and interpretation are unchanged from the linear case. So the only restriction on functional form imposed by the sorting model is that the joint effect of education and productivity on wages is additively separable. The impact of relaxing the logistic assumption is discussed in Section 3.5.

The sorting process makes two strong assumptions - that workers sort on a one-dimensional characteristic, and that they sort perfectly on that characteristic. The assumption that workers sort on a one-dimensional variable provides tractability, but can easily be generalized. Suppose instead that each worker is characterized by two characteristics, a productivity term p_i and a social characteristics term x_i. which takes on values from a discrete set. The social characteristics term can include things like race, ethnic origin, family size, preference for living in cities or suburbs, or any other non-income variables that affect one's choice of residence. Modify the sorting rule so that workers sort first on x_i, then sort on p_i within groups. Note that p_i and x_i may be correlated, so that any effect of race on income can be accounted for directly by p_i. Again, the analysis and interpretation is unchanged by this generalization.

The assumption that sorting is perfect (with enough neighborhoods) is also easily generalized, but requires a slight reinterpretation of results. With either perfect or imperfect sorting, we can use Equation (19) to estimate the distribution of neighborhood average p_i across neighborhoods. With perfect sorting every resident of a given neighborhood has identical p_i, so this distribution is identical to the distribution of p_i across individuals. Because the estimates here conflict with the characteristics of independent estimates of this distribution, we reject the model. With imperfect sorting, the distribution of p_i across neighborhoods and across individuals are not in general identical. For example, if neighborhoods are perfectly integrated by income, the distribution across neighborhoods is degenerate even if there is substantial heterogeneity across individuals. However, the difference between these two distributions will be small if the sorting is almost perfect, and the relative importance of the sorting effect will be lower if sorting is far from perfect. Although quantifying this trade off is difficult, explaining the nonlinearity of the relationship between neighborhood unemployment and neighborhood education would seem to require a high degree of sorting.

Finally, there is the "no social interaction effect" assumption. As shown in Section 2.1, a model with an endogenous neighborhood effect can generate discontinuity in the relationship between neighborhood average education and neighborhood unemployment rate without any sorting effect. If such an endogenous neighborhood effect is added to the sorting model, the nonlinearity we observe in the data can be explained without requiring a negatively skewed productivity distribution.

3.5 Robustness to alternative functional forms

Equations (18) and (19), combined with the assumption that the distribution of the error terms $F_\epsilon(.)$ is logistic, pin down the distribution $F_p(.)$ of the productivity term that fits the data. One might wonder how sensitive the results in Section 3 are to the assumption that the individual-level idiosyncratic effect has the logistic distribution.

Judging the robustness of the results to a particular alternative distribution for that term is simply a matter of applying equations (18) and (19) to a different $F_\epsilon(.)$. In results which are not reported in detail here, I find that replacing the assumption of a logistic distribution with that of a normal distribution has virtually no effect on any of the results.

It is also possible to approach this question from the other direction. Suppose that we assume a particular distribution for productivity. Then there must be a distribution of the error term which solves equations (18) and (19) for the data. In fact, as long as $E(\bar{u}_n|\bar{c}_n)$ is strictly decreasing in $\bar{c}_n$, there are many such distributions, and the data can be used to characterize this set of distributions.

Suppose that the distribution of productivity is positively skewed, as would be predicted on the basis of microeconomic studies. In particular, suppose that its distribution is identical to that of $\bar{c}_n$. This will imply that the relationship $p(\bar{c}_n)$ will be exactly linear; small differences between the shapes of the two distributions will produce small deviations from linearity. In principle, it is possible to now apply equations (18) and (19) to solve for $F_\epsilon(.)$. However, doing so nonparametrically is difficult, so I make a slight approximation. I assume that the probability of unemployment takes the following form:

$$\Pr(u_i = 1|c_i, p_i, \bar{c}_n, \bar{\theta}_n) = F_\epsilon\left(-(\beta_{int} + \beta_c c_i + \beta_p p_i)\right) + \bar{\theta}_n \tag{23}$$

This formulation differs slightly from equation (14) in that the neighborhood-level shock $\bar{\theta}_n$ is outside of $F_\epsilon(.)$. In this case we have:

$$F_\epsilon(-X) = E(\bar{u}_n|\bar{c}_n = X) \tag{24}$$

In other words, "flipping" the nonparametric regressions in Figures 5-8 will give a portion of the CDF $F_\epsilon(.)$ (as always, up to a linear transformation). Figure 15 shows the results for Chicago.

These implied (partially censored) CDF's show some unusual characteristics. The CDF basically looks like a pair of straight lines which meet at a kink. To emphasize this pattern, the graph also shows a line indicating the normal CDF which is closest to this CDF (in Euclidean distance). In summary, the results in this paper are moderately robust to the distribution assumed for the individual idiosyncratic term. Using a normal rather than a logistic changes nothing. However, there are distributions that do not look like a normal or logistic distribution that would generate more support for the sorting model.

4 Conclusion

The results in this paper suggest that epidemics and threshold effects in neighborhoods are not just a theoretical possibility. Neighborhoods in which fewer than twenty percent of over-25 residents has a Bachelor's degree experience much larger unemployment rates, a stylized fact that 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, the explanation implies an implausible wage distribution.

The paper also shows that structural models of sorting 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.

While the empirical relevance of thresholds is a potentially important question, many related questions remain unanswered. In particular, the approach used here is unable to quantify the strength of the social interaction effect, and thus the impact on employment status a change in environment would produce. As a result, the quantitative impact of moving a worker to a neighborhood with higher average educational attainment is uncertain. Experimental studies such as the Moving To Opportunity program analyzed by Katz, Kling, and Liebman [1999] may shed more light on this subject. In addition, the analysis takes labor demand as given. Even if the impact of moving a single worker were known, general equilibrium effects further complicate the use of the results here to suggest large-scale policies to increase employment in poor neighborhoods. Finally, the model here leaves unspecified the whether the social interaction effects arise from job networking, attitude formation, or other sources. While all of these issues remain for future research, the possibility of an threshold effect which can be exploited by public policy is promising.

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. *Review of Economic Studies*, 2000. Forthcoming.

[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] Dennis Epple and Holger Sieg. Estimating equilibrium models of local jurisdictions. *Journal of Political Economy*, 107:645–680, 1999.

[6] 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.

[7] Malcolm Gladwell. *The tipping point: How little things can make a big difference.* Little, Brown, 2000.

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

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

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

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

[12] 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.

[13] Lawrence F. Katz, Jeffrey R. Kling, and Jeffrey B. Liebman. Moving to Opportunity in Boston: Early impacts of a housing mobility program. Working paper, Harvard University, 1999.

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

[15] 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.

[16] Charles F. Manski. *Identification Problems in the Social Sciences.* Harvard University Press, 1994.

[17] Charles F. Manski. Economic analysis of social interactions, in theory and practice. *Journal of Economic Perspectives*, 2000. Forthcoming.

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

[19] 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.

[20] Giorgio Topa. Social interactions, local spillovers, and unemployment. *Review of Economic Studies*, 2000. Forthcoming.

[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.

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.

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 $E(\bar{u}_n|\bar{c}_n)$

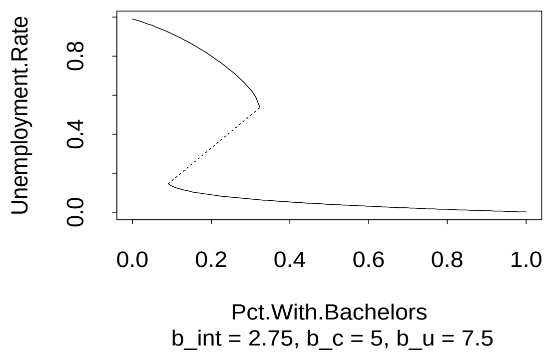

Figure 1: Equilibria of the model with social interactions. The dotted line denotes unstable equilibria. Parameter values are $\beta_{int} = 2.75$, $\beta_c = 5$, and $\beta_u = 7.5$

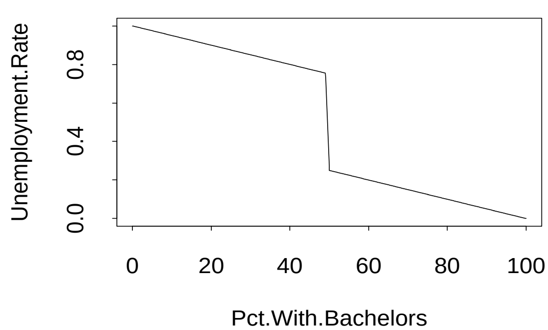

Figure 2: An example of a threshold relationship between two variables.

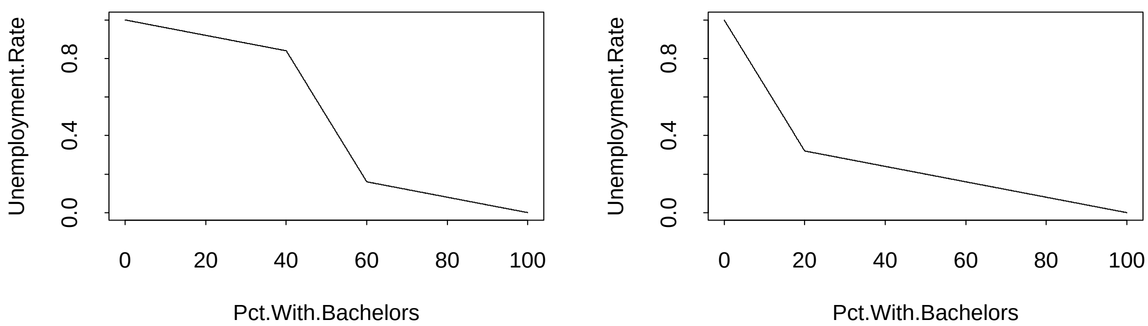

Figure 3: Examples of regression relationships which provide evidence for thresholds.

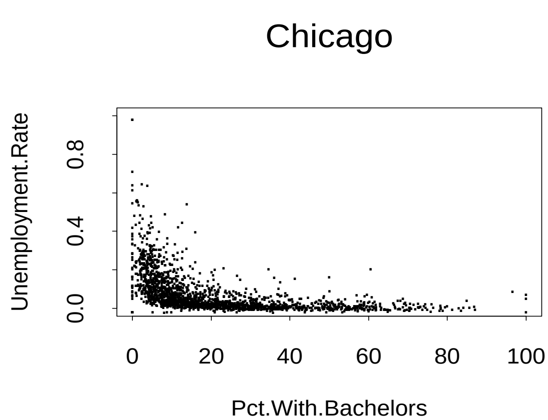

Figure 4: Scatter plot of neighborhood unemployment versus neighborhood education in Chicago CMSA.

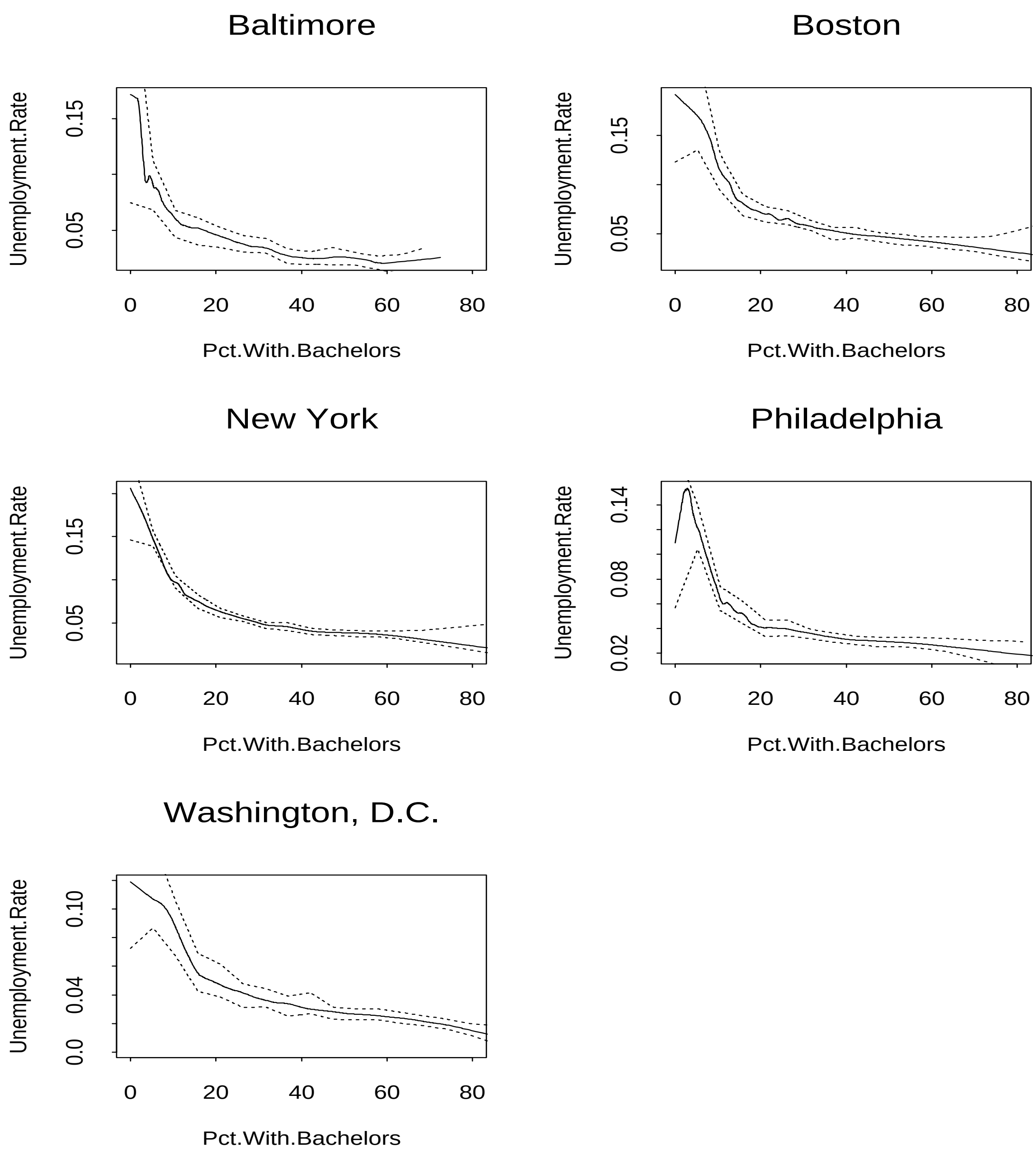

Figure 5: Nonparametric regressions for northeastern cities.

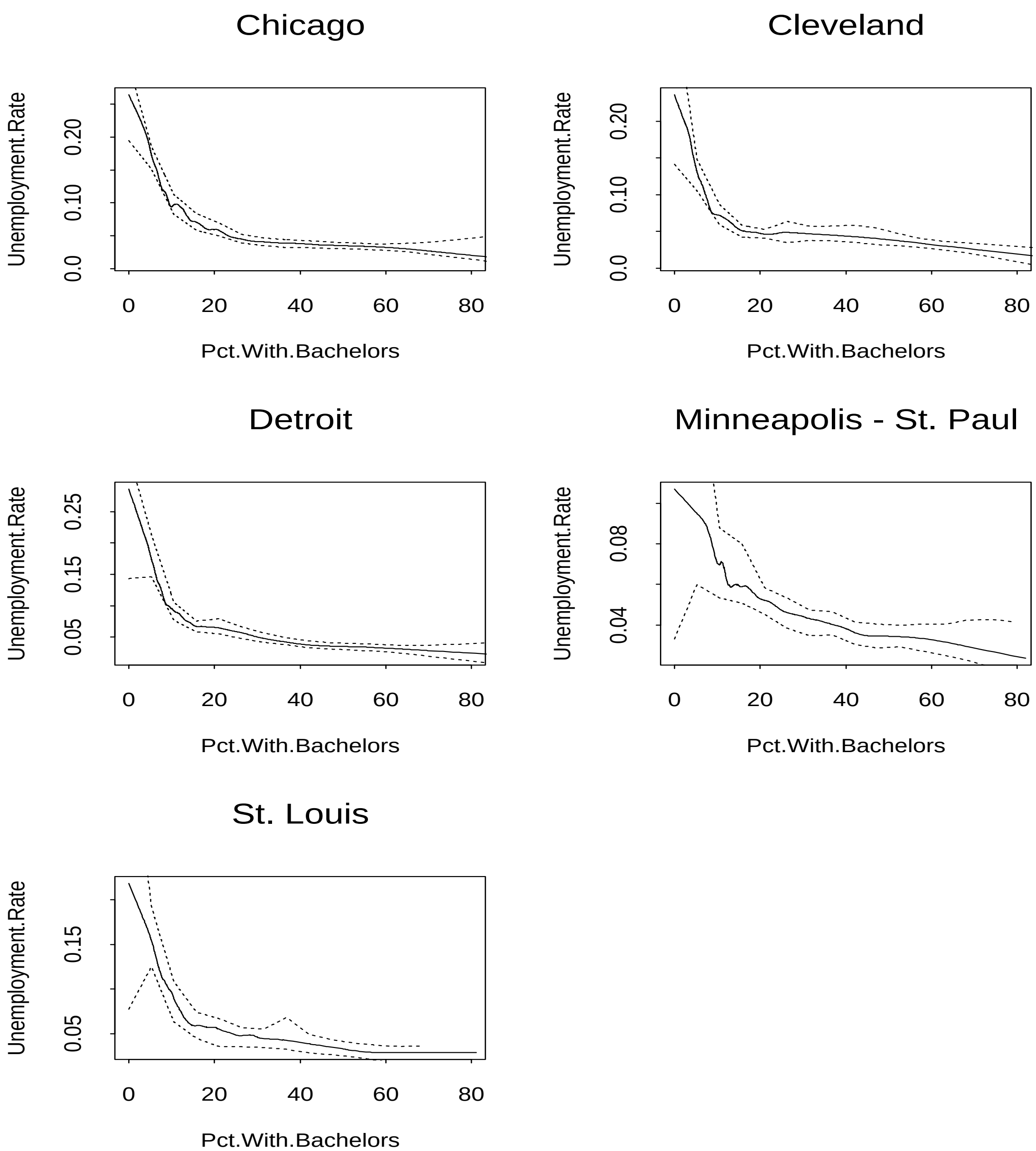

Figure 6: Nonparametric regressions for midwestern cities.

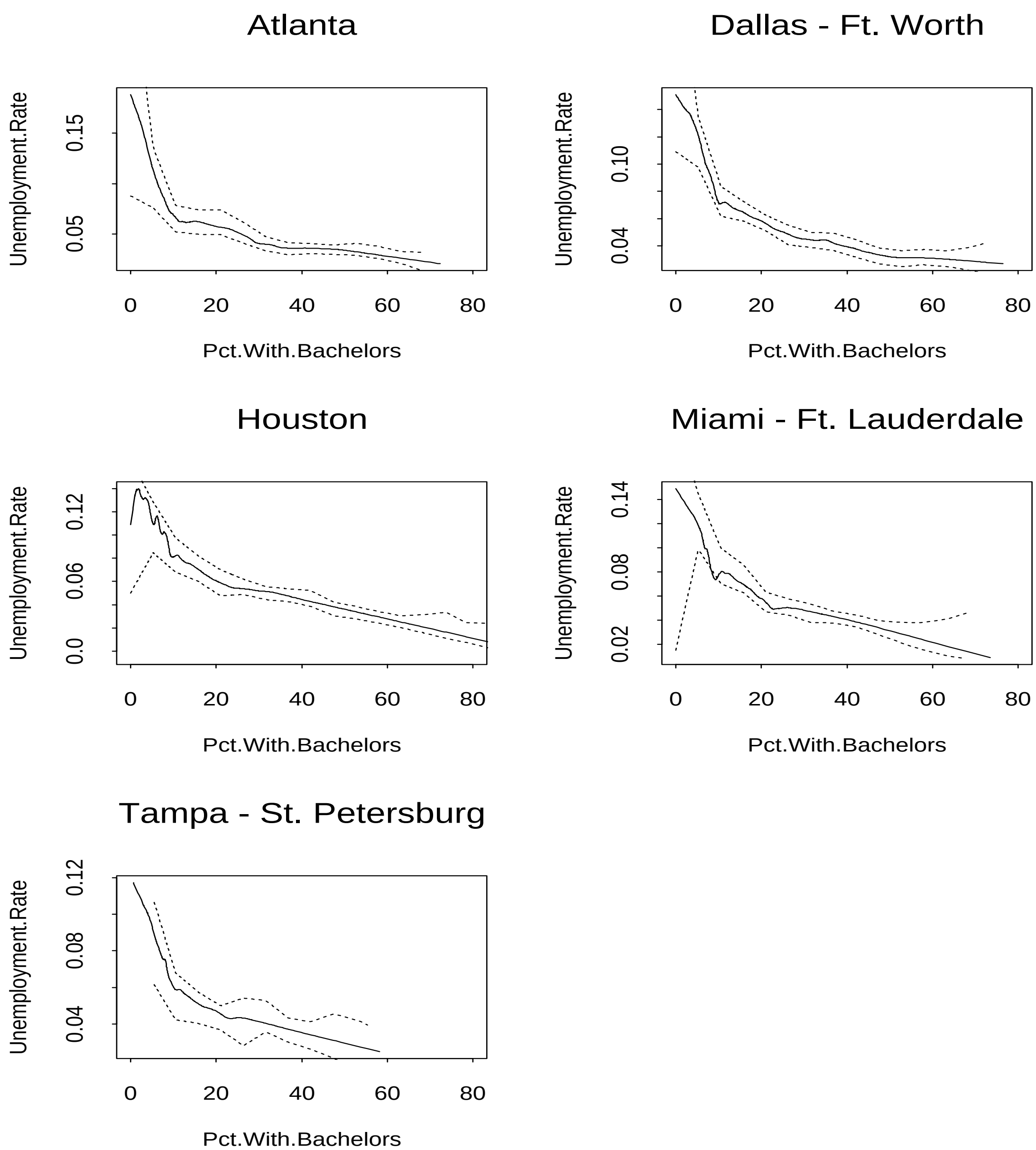

Figure 7: Nonparametric regressions for southern cities.

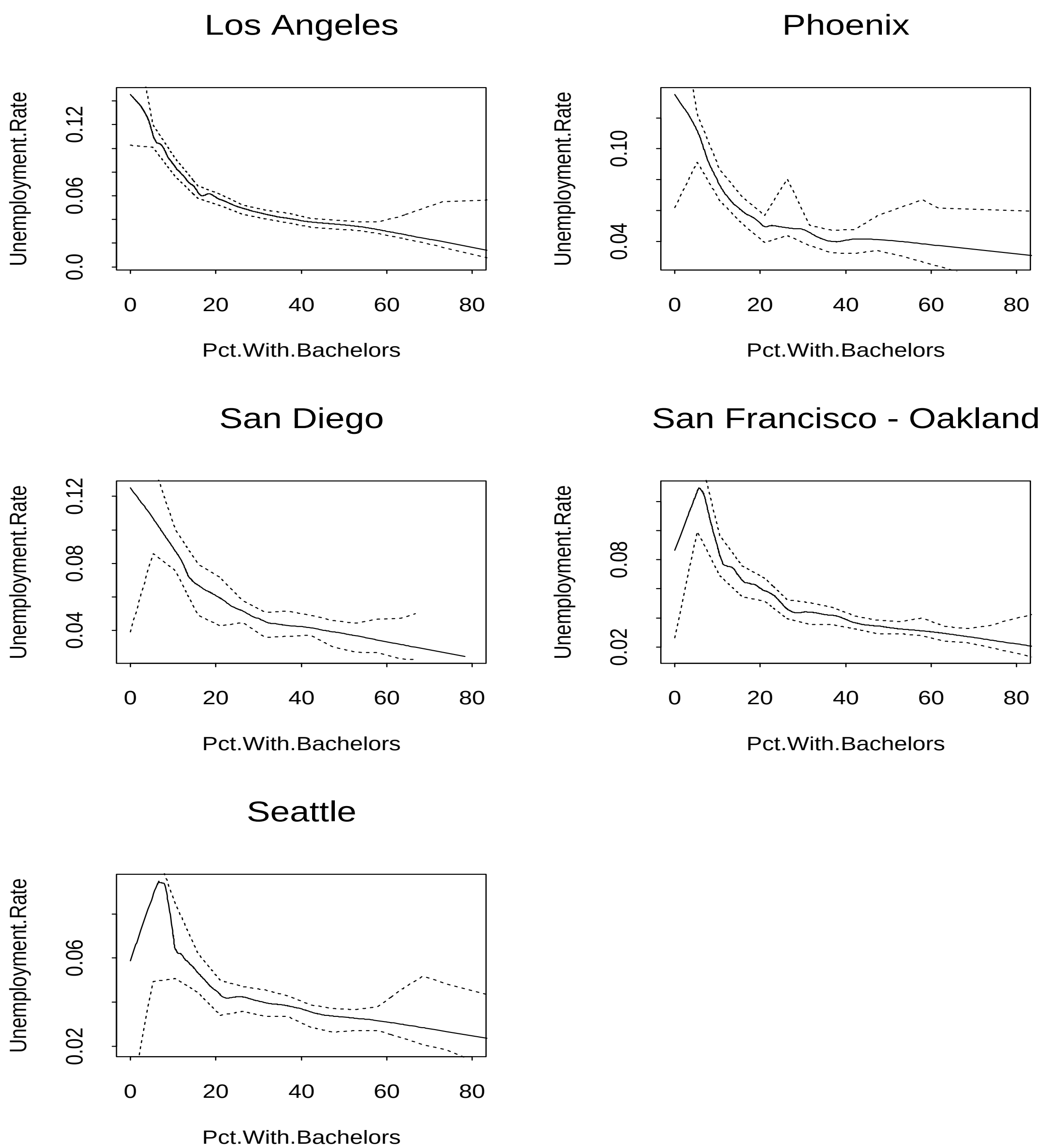

Figure 8: Nonparametric regressions for western cities.

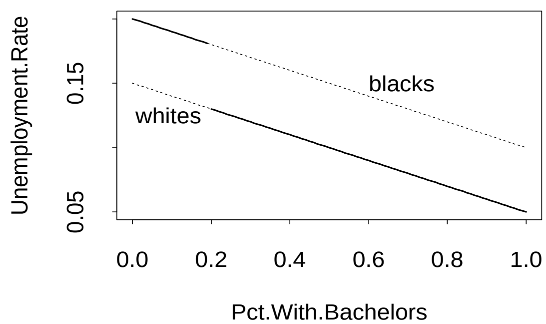

Figure 9: Example of spurious threshold due to ethnic group differences in unemployment rates.

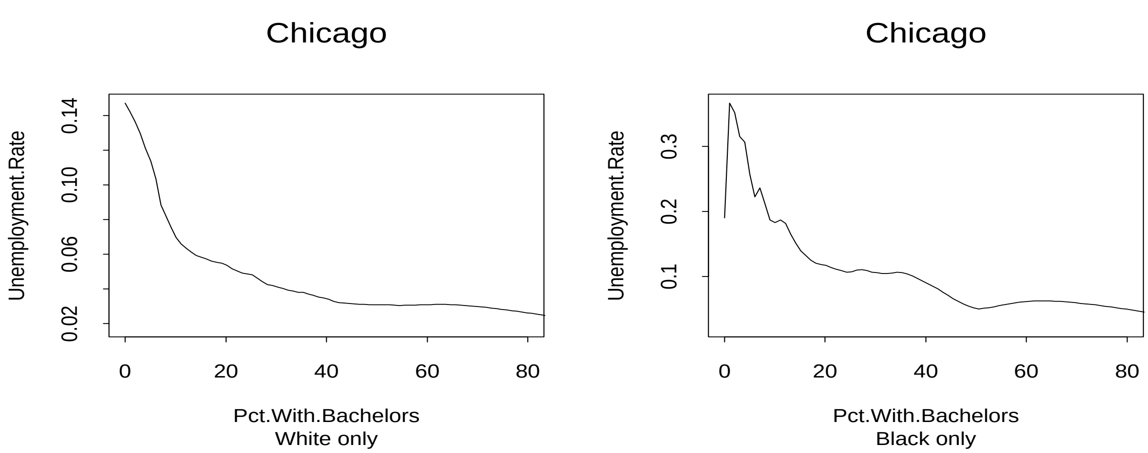

Figure 10: Nonparametric regressions by racial category. First graph shows estimated regression for whites, second shows regression for blacks.

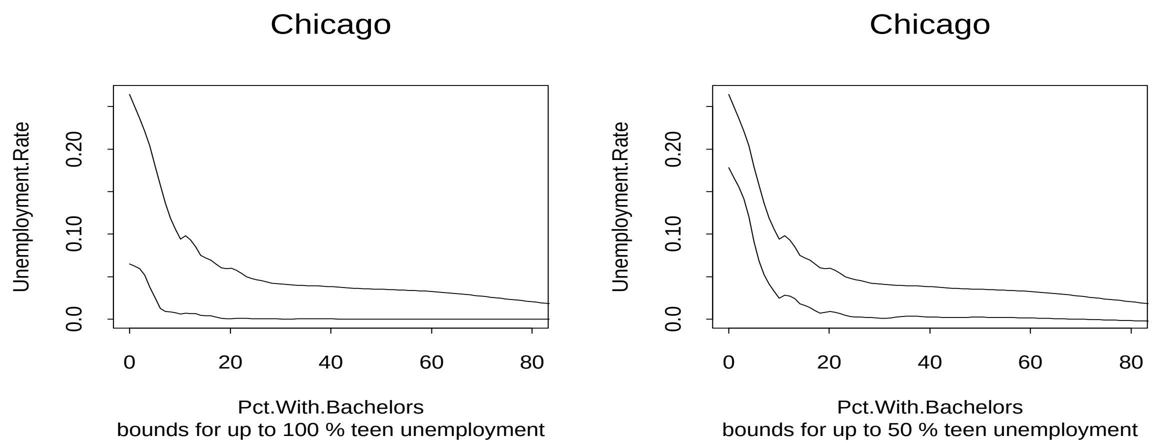

Figure 11: Bounds on regression function, Chicago CMSA. The first graph shows the worst-case bounds of 100% youth unemployment, the second shows an upper bound of 50% youth unemployment.

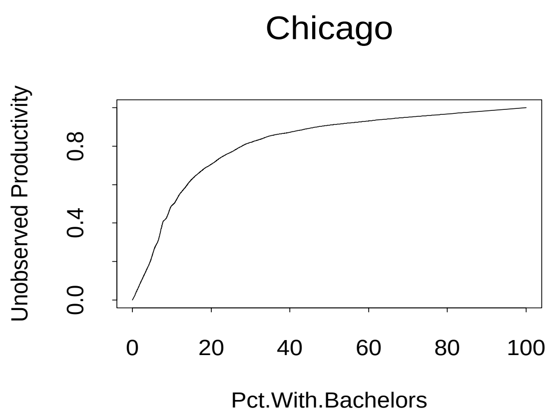

Figure 12: Estimated $p(\bar{c}_n)$ for Chicago. β_{int} and β_p set so that range of $p(\bar{c}_n)$ is $[0, 1]$.

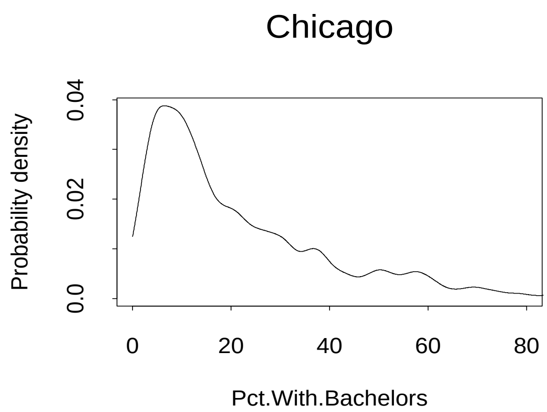

Figure 13: Distribution of $\bar{c}_n$ across neighborhoods in Chicago.

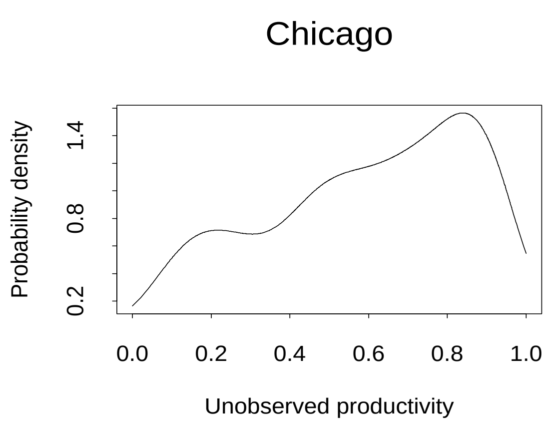

Figure 14: Distribution of p_i across individuals in Chicago, for estimated $p\left(\bar{c}_n\right)$.

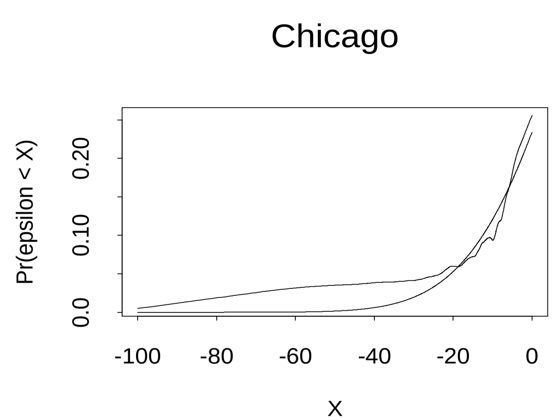

Figure 15: Estimated CDF of ϵ_i under assumption of a linear $p\left(\bar{c}_n\right)$. The smooth curve is the closest (by Euclidean distance) normal CDF.

A Notes for referees

A.1 The college wage premium

In Section 3, I assume that the parameter β_c, 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_c.

The starting point is equation (16) which is rewritten here:

$$\bar{u}_n = \bar{c}_n * F_\epsilon\left(-(\beta_{int} + \beta_c + \beta_p p\left(\bar{c}_n\right) + \bar{\theta}_n)\right) + (1 - \bar{c}_n) * F_\epsilon\left(-(\beta_{int} + \beta_p p\left(\bar{c}_n\right) + \bar{\theta}_n)\right)$$

Instead of assuming that $\beta_c = 0$, assume any positive value. Define $\bar{z}_n$ for a particular neighborhood as the solution to:

$$\bar{u}_n - \bar{c}_n * F_\epsilon\left(-(\beta_c + \bar{z}_n)\right) - (1 - \bar{c}_n) * F_\epsilon\left(-\bar{z}_n\right) = 0 \tag{25}$$

Because the left side of equation (25) is a monotonic function which passes through zero, $\bar{z}_n$ can be found for each neighborhood as a function of the observable variables $\bar{u}_n$ and $\bar{c}_n$. From equation (16), we also have:

$$\bar{z}_n = \beta_{int} + \beta_p p\left(\bar{c}_n\right) + \bar{\theta}_n \tag{26}$$

Because $E(\bar{\theta}_n|\bar{c}_n) = 0$, the function $p\left(\bar{c}_n\right)$ can thus be estimated (up to the usual linear transformation) using:

$$p\left(\bar{c}_n\right) = E(\bar{z}_n|\bar{c}_n) \tag{27}$$

So it is possible to determine the $p\left(\bar{c}_n\right)$ function implied by the data for a given β_c.

Figure 16 shows the estimated $p\left(\bar{c}_n\right)$ function for several "reasonable" values of β_c. As the figure shows, the estimated $p\left(\bar{c}_n\right)$ looks essentially the same for the values $\beta_c = 0, 0.1, 0.5$, and becomes non-monotonic for $\beta_c = 2$ (and higher). In other words, over a reasonable range (i.e., one that does not imply either a negative college wage premium or a non-monotonic $p\left(\bar{c}_n\right)$), the estimated $p\left(\bar{c}_n\right)$ is not sensitive to the value of β_c.

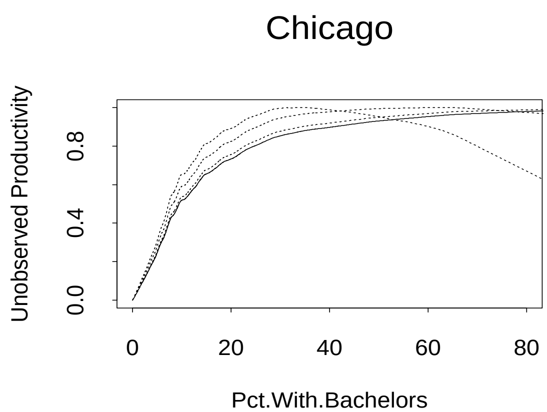

Figure 16: Estimated $p\left(\bar{c}_n\right)$ for Chicago CMSA, assuming values for β_c of $\{0, 0.1, 0.5, 2\}$.