University of Connecticut

Department of Economics Working Paper Series

Identifying Individual and Group Effects in the Presence of Sorting: A Neighborhood Effects Application

Patrick Bayer
Yale University

Stephen L. Ross
University of Connecticut

Working Paper 2006-13

April 2006

341 Mansfield Road, Unit 1063
Storrs, CT 06269–1063
Phone: (860) 486–3022
Fax: (860) 486–4463
http://www.econ.uconn.edu/

Abstract

Researchers have long recognized that the non-random sorting of individuals into groups generates correlation between individual and group attributes that is likely to bias naïve estimates of both individual and group effects. This paper proposes a non-parametric strategy for identifying these effects in a model that allows for both individual and group unobservables, applying this strategy to the estimation of neighborhood effects on labor market outcomes. The first part of this strategy is guided by a robust feature of the equilibrium in the canonical vertical sorting model of Epple and Platt (1998), that there is a monotonic relationship between neighborhood housing prices and neighborhood quality. This implies that under certain conditions a non- parametric function of neighborhood housing prices serves as a suitable control function for the neighborhood unobservable in the labor market outcome regression. The second part of the proposed strategy uses aggregation to develop suitable instruments for both exogenous and endogenous group attributes. Instrumenting for each individuals observed neighborhood attributes with the average neighborhood attributes of a set of observationally identical individuals eliminates the portion of the variation in neighborhood attributes due to sorting on unobserved individual attributes. The neighborhood effects application is based on confidential microdata from the 1990 Decennial Census for the Boston MSA. The results imply that the direct effects of geographic proximity to jobs, neighborhood poverty rates, and average neighborhood education are substantially larger than the conditional correlations identified using OLS, although the net effect of neighborhood quality on labor market outcomes remains small. These findings are robust across a wide variety of specifications and robustness checks.

The authors are grateful to the Department of Housing and Urban Development, Federal Reserve Bank of New York, and the Center for Real Estate and Urban Economic Studies at the University of Connecticut for financial support. We are especially thankful to Steven Durlauf for extensive comments on an earlier draft and to William Brock, Fernando Ferreira, Albert Saiz, Holger Sieg, Todd Sinai and seminar participants at UC Berkeley, NYU, Penn-Wharton, and Wisconsin for many helpful comments and suggestions. The research in this paper was conducted while the authors were Special Sworn Status researchers of the U.S. Census Bureau at the Boston Research Data Center (BRDC). The research results and conclusions expressed are those of the authors and do not necessarily reflect the views of the U.S. Census Bureau, the Department of Housing and Urban Development or any other agency of the U.S. Government. This paper was screened to insure that no confidential data are revealed.

1 Introduction

Economists often examine economic behavior and outcomes in empirical settings in which individuals have non-randomly sorted into groups. Examples include schools, residential neighborhoods, occupations, interpersonal relationships, correctional and treatment programs. In attempting to separately identify the impact of individual versus group attributes on individual outcomes, researchers have long recognized that this non-random sorting generates correlation between individual and group attributes some of which are likely to be unobserved. Because the associated biases are likely to be especially severe in the estimation of social interactions (peer effects) among individuals in the same reference group, the identification problems created by endogenous group formation have received extensive attention in that literature (Manski (1993), Moffitt (2001), Brock and Durlauf (2001)).

In this paper, we consider the general problem of identifying the effect of individual and group attributes on individual outcomes in a model that allows *for both individual and group unobservables*. When both types of unobservables are present, the identification problem created by non-random sorting is especially severe because, as in Epple (1987), any variable that affects sorting over groups will generally, by construction, be correlated with either the individual or group unobservable (or both).[2] As a result, many common empirical strategies that are designed to deal with the presence of one type of unobservable often neglect the presence of the other, thereby either not addressing or possibly even exacerbating the corresponding biases.

To see this, consider a specific estimation technique that is used commonly in the literature: the inclusion of group fixed effects in the individual outcome regression. While it is often argued that this approach eliminates any biases resulting from across-group sorting, thereby providing unbiased estimates of the effect of observable individual attributes, this is not generally the case. The problem is that non-random sorting generally induces correlation between observed and unobserved individual attributes within groups. In a selective hiring or matching process (e.g., teachers to schools, individuals to colleges, doctors to residency programs, individuals to occupations), for example, the total ability of individuals that sort into the same group is likely to be comparable leading to negative within-group correlation between observed and unobserved measures of individual ability, even if these measures are uncorrelated in the population. In such

[2] As we discuss in greater detail below, the identification problem induced by the presence of two types of unobservables in our model bears a close resemblance to the identification problem in hedonic models described by Epple (1987). In that case, Epple showed that the equilibrium matching of suppliers to consumers along the hedonic gradient ensured that unobservable attributes from both the demand- and supply-side of the problem enter the estimating equation, thereby making classic instruments derived from the opposite side of the market inappropriate and the identification essentially impossible with observational data.

cases, the inclusion of group fixed effects would generally lead to an attenuation bias in the estimated effect of observed individual attributes.[3] More generally, the use of group fixed effects does nothing to specifically address biases that arise when sorting is driven in part by unobserved individual attributes.

To address the identification problem induced by the presence of both individual and group unobservables in an outcome regression, we offer a non-parametric solution that is grounded in the canonical vertical model of sorting developed in Epple and Platt (1998) and Epple and Sieg (1999).[4] To make the discussion of the identification problem and our proposed empirical methodology concrete, we consider a specific application: the estimation of effect of neighborhood attributes on labor market outcomes.

We begin the paper by using the structure of the sorting equilibrium in the Epple-Platt-Sieg (EPS) model to highlight the correlations that are induced when individuals sort across neighborhoods on the basis of both individual and neighborhood unobservables. We then present our identification strategy, which consists of two distinct parts. We begin by exploiting a robust implication of the sorting equilibrium in the EPS model, namely that there is a monotonic relationship between neighborhood housing prices and neighborhood quality. Under conditions that we make explicit below, this implies that a non-parametric function of neighborhood housing prices serves as a suitable control function for the neighborhood unobservable in the labor market outcome regression.[5] [6] By including this control function in the labor market outcome regression, we eliminate the correlation between the group component of the error term and the individual attributes included in the regression. This strategy implements the general observation made in Brock and Durlauf (2006) that a control function can be used to deal with the group unobservable in an individual outcome equation.[7]

---

[3] It is not uncommon to see groups fixed effects used in estimation for environments where sorting across groups is expected. For example, the inclusion of occupation fixed effects in wage regressions, school fixed effects in models of teacher productivity, or neighborhood fixed effects in models of housing prices.

[4] This model and its predecessors has been used or extended in theoretical settings by Epple, Filimon, and Romer, (1984, 1993), Epple and Romer (1991), Epple and Romano (1999), Fernandez and Rogerson (1996, 1998) and Benabou (1993, 1996) and applied in empirical settings by Epple, Romer, and Sieg (2001), Sieg, Smith, Banzhaf, and Walsh (2004), Walsh (2005) among others. See Ross and Yinger (1999) for a review of papers that apply this sorting model within local public finance.

[5] Note that monotonic relationship may not hold explicitly once one allows for horizontal sorting as in the models developed by Nechyba (1997, 1999), Bayer, McMillan, and Rueben (2005), Bayer and Timmins (2005, 2006), Ferreira (2003) and Ferryera (2003).

[6] In addition to the residential sorting context, this solution should be applicable in any setting where the price of entry into a group is available (e.g., wages, college tuition) or where groups can be quality rank-ordered in some way.

[7] Ioannides and Zabel (2004) use such a control function in their work on housing demand. It is the specific idea to use neighborhood housing price as a control function for unobserved neighborhood quality in an individual outcome equation that is new here. Also, note that the use of price as a control function

The second part of our strategy is designed to address the likely correlation of unobserved individual attributes with observed group attributes (including the housing price control function) and follows a more traditional IV approach. To break this correlation, we assign each individual in the sample to a cell based on her observable characteristics and instrument for each individual's own neighborhood attributes with the average neighborhood attributes of those individuals in the same cell. Averaging neighborhood attributes over all observationally equivalent individuals removes any idiosyncratic portion of the sorting of individuals into neighborhoods associated with an individual's unobservable attributes. Notice also, that this approach amounts to using a fully non-parametric sorting model to predict each individual's neighborhood attributes given her observable characteristics. Brock and Durlauf (2001, 2002, 2005) have recommended the use of such non-linearities arising from discrete choices for identification in models of social interactions.[8] [9]

For our neighborhood effects application, we use the confidential Long Form data from the 1990 Decennial Census for the Boston Metropolitan Statistical Area. In examining the impact of neighborhoods on labor market outcomes, we focus on the influence of spatial access to jobs and neighborhood socioeconomic characteristics on individual labor market outcomes. These neighborhood attributes have received a great deal of attention in the previous literature. We estimate models for six different labor market outcomes, a number of subsamples based on education, gender, and family structure, and a variety of empirical specifications designed to isolate the impact of each of the three parts of our proposed identification strategy.

Our results imply that the direct impact of geographic proximity to jobs, neighborhood poverty rates, and college-educated neighbors is substantially larger than the conditional correlations identified using OLS. These findings are robust across a wide variety of specifications and robustness checks. Interestingly, while geographic proximity and neighborhood poverty rates have the anticipated positive and negative impacts on labor market outcomes respectively, exposure to college-educated neighbors also has a significant negative effect. We discuss potential explanations for this finding below. Thus, taken together, our results imply that the relationship between neighborhood attributes and labor market outcomes is quite complex and as a whole our results are consistent with small and even negative net effects of

---

has also appeared in the differentiated products demand literature in Petrin and Train (2005), although it is used there primarily as a computational tool in a standard instrumental variables context.

[8] Also see Bayer and Timmins (2006) and in the context of the identification of hedonic models by Ekeland, Heckman, and Nesheim (2004), Bajari and Benkhard (2005) and Bajari and Kahn (2005).

[9] In addition to the two-part strategy outlined here, we also address additional issues related to neighborhood attributes endogenously determined by the sorting process (e.g., neighborhood socioeconomic characteristics) below.

improving neighborhood 'quality' on the labor market outcomes. The finding of small net effects of neighborhood is primarily driven by the negative effect of college-educated neighbors.

The remainder of the paper is organized as follows. Section 2 provides a broader review of the neighborhood effects literature. Section 3 presents a simple version of the canonical vertical sorting model of Epple and Platt (1998) and Epple and Sieg (1999) and examines the resulting biases in ordinary least squares analyses of the effect of individual and group attributes on individual outcomes. Section 4 presents our three-part estimation strategy for obtaining consistent estimates in the presence of both individual and group unobserved attributes. Section 5 discusses the data, sample, and specification of variables used to describe households and neighborhoods. Section 6 presents the results and Section 7 concludes.

2. Neighborhood Effects and Labor Market Outcomes – Previous Literature

For the purposes of our analysis two aspects of the previous neighborhood effects literature are pertinent. First, as we seek to offer a general solution to a core identification problem in the neighborhood effects literature, we begin by reviewing other recent approaches to the problem, noting their strengths and limitations. Second, we then discuss the previous empirical findings in the literature that relate most directly to our application: the effect of neighborhood on labor market outcomes.

Identifying Neighborhood Effects. The study of the identification of neighborhood effects is a difficult problem without a completely general solution. An important line of recent research seeks to identify neighborhood effects by isolating a random component of neighborhood choice induced by special social experiments. Popkin et al. (1993) pioneered this approach using data from the Gautreaux Program conducted in Chicago in the late 1970's, which gave housing vouchers to eligible black families in public housing as part of a court-imposed public housing de-segregation effort. Similarly, Oreopolous (2003) and Jacob (2005) study the impact of re-locations arising from administrative assignment to public housing projects in Toronto and from the demolition of the public housing projects in Chicago, respectively. Most notably, Katz et. al. (2001) and Ludwig et al. (2001) have used the randomized housing voucher allocation associated with the Moving To Opportunity demonstration (MTO) to examine the impact of re-location to neighborhoods with much lower poverty rates on a very wide set of individual behavioral outcomes including health, labor market activity, crime, education, and more. Especially in the case of MTO, the advantages of this approach are clear – the randomization inherent in the program design ensures a clean comparison of treatment and proper control groups.

There are, however, important limitations in the extent to which the treatment effects identified through re-location are informative about the nature of general forms of neighborhood effects *per se*. First, individuals studied must be eligible for a re-location program in the first place; this typically implies that the resulting sample is special (i.e. so as to be a resident in public housing) and may not be as sensitive to neighborhood effects as other individuals. Second, the experimental design involves re-location to new neighborhoods that are, by design, very different from baseline neighborhoods; this implies that the identified treatment effect measures the impact of re-locating to a neighborhood where individuals initially have few social contacts and where the individuals studied may be very different than the average resident of the new neighborhood. In this way, the treatment effects identified with this design are necessarily a composite of several factors related to significant *changes* in neighborhoods that are not easily disentangled (see Moffitt (2001) for a detailed discussion).

A second broad approach seeks to deal with the difficulties induced by correlation in unobserved attributes at the neighborhood level by aggregating to a higher level of geography. Evans, Oates, and Schwab (1992), Cutler and Glaeser (1997), Ross (1998), Weinberg (2000, 2004), Ross and Zenou (2004), and Card and Rothstein (2005) identify the effect of location on outcomes using cross-metropolitan variation. For example, Cutler and Glaeser (1997) analyze the impact of segregation within a metropolitan area on a variety of outcomes including education, labor market activity, and teenage fertility, and Evans, Oates and Schwab use metropolitan area poverty rates as an instrument for neighborhood level poverty. Again, the advantages of this approach are clear – aggregation certainly eliminates the problem of correlation in unobservables among neighbors (although potential correlation in unobservables at the metropolitan level becomes an issue). The effects identified through aggregation, however, include not only the average *neighborhood* effects operating in a metropolitan area but also any broader consequences of living in a segregated or high poverty metropolitan area.[10] Thus, the strict interpretation of the estimated effects as neighborhood effects requires the assumption that metropolitan segregation does not directly affect outcomes.[11]

A third approach is offered in Bayer, Ross, and Topa (2005), which uses detailed Census microdata to isolate block-level variation in the characteristics of neighbors within narrowly-defined neighborhoods. The key identifying assumption underlying this design is that there is no

[10] More residentially segregated metropolitan areas might be associated, for example, with increased racial taste-based discrimination in the labor market, in the application of criminal justice, etc. due to decreased levels of regular inter-racial contact in residential neighborhoods.
[11] It is important to point out that Cutler and Glaeser (1997) do not claim that the effects identified in their analysis are strictly neighborhood effects.

block-level correlation in unobserved attributes within block groups, due perhaps to the thinness of the housing market (i.e., that it is difficult to select the particular block that one would like to live on). This approach identifies the effect of neighborhood by conditioning on the effect of location over a broader geographic range (block group). Accordingly, the methodology is designed to capture very local social interactions whose influence decays very quickly with distance. This approach will not capture effects that arise over broader areas, such as the influence of employment access or crime rates on employment outcomes.

The Epple-Platt-Sieg style model that we will discuss below provides a useful framework for understanding these studies. The model below contains two sources of error over which individuals sort: an individual specific error and a neighborhood specific error. In this context, traditional instrumental variables analysis often fails. For example, an instrument for an individual's neighborhood attributes must be correlated with neighborhood choice and yet uncorrelated with either the individual or place unobservable. However, these three requirements represent a contradiction in a sorting equilibrium. Any variable that is uncorrelated with the individual unobservable will only be correlated with neighborhood choice if it influences sorting over place unobservables, but of course this contradicts the assumption that the instrument is uncorrelated with the place unobservable.[12] In experimental studies, residential location is changed based on a randomly assigned experimental treatment. While in the studies that use across metropolitan variation or in the study that uses within block group variation, the implied assumption is that the factors that influence across metropolitan sorting or within block group sorting are idiosyncratic and orthogonal to the individual and neighborhood unobservables that influence labor market outcomes.

All of the above studies use an empirical design intended to provide a variable that is correlated with an individual's exposure to location attributes for reasons that are independent of the individual's sorting behavior (experimental treatment, exogenous residence in metropolitan areas, or random sorting arising from a thin housing market), but these designs in turn limit the researchers' ability to isolate and identify the effects of various neighborhood factors on current residents. By systematically addressing the sources of sorting bias in a population of current neighborhood residents, we hope to provide a more detailed and complete picture of the influence neighborhood on resident outcomes albeit with newly imposed assumptions concerning the structure of the underlying sorting problem. Specifically, unlike most of the studies above, this study is intended to capture the overall effect of neighborhood variables on a representative

[12]As mentioned earlier, this identification problem is comparable to Epple's (1987) analysis of identification in estimating hedonic models.

population in their equilibrium locations. Moreover, our approach generates a substantial amount of identifying information for relatively large population-based samples. Therefore, unlike the studies above that rely on across-metropolitan variation or small experimental samples and so are only able to examine one or two variables of interest, our study is able to examine the effect of a larger variety of neighborhood variables on individual outcomes.

The Effect of Neighborhood on Labor Market Outcomes. A wide array of studies have documented the relationship between various aspects of the neighborhood environment and employment outcomes. The spatial mismatch hypothesis, first proposed by Kain (1968), has spawned innumerable studies that find that job access is positively correlated with employment and/or labor market earnings. Ihlanfeldt and Sjoquist (1990) and Raphael (1998), for example, find that youth residing far from suburban areas where low skill jobs tend to be locatedand where new jobs tend to be created had worse employment outcomes. Other research has centered on the impact of the characteristics and behavior of neighbors on labor market outcomes. Case and Katz (1991), for example, find a correlation between youth idleness and the idleness of neighbors, while O'Regan and Quigley (1998) find that youth are more likely to be high school dropouts and unemployed when they reside in high poverty neighborhoods and Weinberg, Reagan and Yankow (2004) find that people who move to neighborhoods with worse attributes have worse employment outcomes. [13]

Many scholars have suggested job market referrals or information networks as an important factor behind such neighborhood effects.[14] Rees and Schultz (1970), Corcoran et al. (1980), Holzer (1988), Blau and Robbins (1990), Blau (1992), Granovetter (1995), Addison and Portugal (2001) and Wahba and Zenou (2003) all document the importance of referrals and other informal hiring channels in the labor market, using both U.S. and non-U.S. data. A number of these studies including Holzer (1988) and Blau and Robbins (1990) find that informal referrals are more productive than more formal methods in terms of job offer and acceptance probabilities. Additional studies including Datcher (1983), Devine and Kiefer (1991), Marmaros and Sacerdote

[13] These papers represent a small sample of very large literatures. For broader surveys of these literatures, see Ihlanfeldt and Sjoquist (1998), Ellen and Turner (1997), and Mayer (1996)

[14] The use of informal channels such as referrals by employers can be rationalized as a means to reduce the uncertainty regarding the quality of a prospective employee. Montgomery (1991) was the first to formally model a labor market in which both formal and informal hiring channels coexist. Focusing more closely on the information exchange among workers, Calvo-Armengol and Jackson (2002) analyze an explicit network model of job search in which agents receive random offers and decide whether to use them themselves or pass them on to their unemployed contacts depending on their own employment status and current wage.

(2002), and Loury (2004) find evidence that use of informal networks increases the quality of the match as captured by job tenure or earnings.[15] [16]

Further, this literature suggests that the effect of referrals varies considerably across different demographic groups. In terms of intensity of usage, workers with less education and located in high poverty rate neighborhoods are more likely to use informal contacts (Elliot, 1999), men use referral networks more intensively than women (Corcoran et al., 1980), and Hispanic men use networks more intensively than non-Hispanic white men (Smith, 2000). The productivity of networks also appear to differ across groups with high success rates observed for men relative to women (Bortnick and Ports, 1992) and blacks relative to whites (Bortnick and Ports, 1992; Korenman and Turner, 1996; Holzer (1987). In addition, Bayer, Ross, and Topa (2004) find that both college educated workers and high school drop-outs benefit less than high school graduates from block level employment referrals.[17] They also find that workers with children of similar age are more likely to successfully share employment referral information, and married women are least likely to successfully share employment referral information with each other.

A relationship between labor market outcomes and neighborhood attributes may exist for a variety of reasons. The most commonly discussed mechanisms in the literatures cited above involve information barriers to job search and the significance of informal job market referrals. Residential locations that are far from employment concentrations or have high concentrations of individuals who are not strongly attached to the labor market may provide job searchers with little opportunities for mentoring or for gathering information concerning potential job openings. On the other hand, a high quality neighborhood may provide the individual with neighborhood amenities that are complementary to leisure or may expose individuals to lower risk of adverse events that influence labor market productivity or behavior. For example, Kling, Liebman, Katz, and Sanbonmatsu (2004) find that moving to a low poverty rate neighborhood improves themental and physical health of housing voucher recipients inthe Moving to Opportunity Study (MTO). In fact, the MTO study findings also suggest that there could be multiple mechanisms at work in the relationship between neighborhood and labor market outcomes. MTO implies

[15] See Elliot (1999) and Loury (2003) for counter examples where the use of informal networks led to lower wages. Of course, the lower wages may be associated with increased match quality on desirable job attributes causing the individual to accept a lower wage as a compensating differential.
[16] See Ionnides and Loury (2004) for a detailed review of this literature.
[17] This finding also is consistent with assortive models of social interactions where non-college graduates use informal networks intensively, but college graduates are not part of that network. See Bertrand, Luttmer, and Mullainathan (2000), Aizer and Currie (2004), Arcidiacono and Vigdor (2004), and Weinberg (2005) for similar examples relating to welfare participation, prenatal care use, social interactions at elite universities, and social interactions among high school students, respectively.

substantial neighborhood effects on health for voucher recipients, but no influence on labor market outcomes while many studies document a positive influence of mental and physical health on labor market outcomes.[18] The results from MTO and studies of health and the labor market can only be consistent if there are other influences of neighborhood poverty among MTO recipients that depress labor market activity.

In order to better understand the complex relationship between neighborhood and outcomes, we first focus on three core variables: employment access drawing on the spatial mismatch literature, percent poverty which is a standard measure of neighborhood quality, and percent of college graduates which was intended to proxy for the density of human capital in the neighborhood, but appears to capture either non-linearities in neighborhood referrals or unobserved neighborhood amenities that are associated with a demand for leisure. In addition, we also extend the model to consider the effect of minority and immigrant population shares.

3. Identifying Individual and Group Effects in the Presence of Sorting

This paper posits a world where neighborhoods generate benefits for individuals that might or might not be reflected in their outcomes and individuals sort across neighborhoods trading off the benefits offered by each neighborhood against the price required for access to that neighborhood. In such a world, attributes of both individuals and neighborhoods that affect the sorting process or outcomes may be unobserved to the econometrician. Most existing research only explicitly considers either the individual or the location unobservables or does not make a clear distinction, and yet the interplay of these two unobservables is crucial in understanding the bias arising in any study of neighborhood effects.

The equation that we are interested in estimating can be written as:

$$(1) \qquad y_{ij} = \beta_1 Z_i + \beta_2 X_j + \omega_i + \xi_j + \varepsilon_{ij}$$

where i indexes individuals, j indexes neighborhoods, y_{ij} is the individual outcome of interest, Z_i, ω_i, are observed and unobserved individual attributes respectively, and X_j, ξ_j are observed and unobserved neighborhood attributes, respectively. While we will explicitly allow for endogenous neighborhood attributes in our discussion of the empirical strategy in Section 4, it is expositionally simpler to consider only exogenous attributes here. Specifically, we assume that

[18] For some recent examples, See Smith (2003, 1999), Case, Lubotsky, Paxson (2002), Ettner, Frank, Kessler (1997).

the covariances between observed and unobserved attributes are equal to zero in the distributions of individuals and neighborhoods:[19]

$$(2)\qquad \begin{array}{ll} (i) & E\lfloor X_j{}'\xi_j \rfloor = 0; \\ (ii) & E[Z_i{}'\omega_i] = 0; \end{array}$$

Even when only exogenous attributes are considered in (1), non-random sorting will generally imply correlation between all individual and neighborhood attributes, thereby creating correlation between any observed attributes and the composite error term in (1). To see why non-random sorting gives rise to such correlations, it is helpful to write down a simple version of the Epple-Platt-Sieg model.

A Simple EPS Model of Residential Sorting. Consider a closed metropolitan area consisting of J neighborhoods with a finite number of houses available in each neighborhood. Exogenous neighborhood attributes X_j and ξ_j are distributed such that (2.i) holds. A population of individuals of total size equal to the total number of house available in the metropolitan area individuals has individual characteristics Z_i and ω_i distributed such that (2.ii) holds. Let individuals sort across neighborhoods trading off between the outcome of interest y_{ij} and the price of entering neighborhood j, p_j. Specifically, write individual utility V from choosing neighborhood j as:

$$(3)\qquad V_{ij} = f\left(y_{ij}, p_j, e_i, \mu_i\right)$$

where e_i indicates an individual's initial financial endowment and μ_i represents an individuals tastes for the outcome influenced by neighborhood choice, *y,* versus all other forms consumption, the price of which is assumed to be independent of the individual's neighborhood choice. Given the structure of equation (1), it will be helpful to characterize the neighborhood contribution to the individual outcome *y* as:

$$(4)\qquad \theta_j = \beta_2 X_j + \xi_j$$

[19] These assumptions of strict exogeneity are standard in any simple regression estimate of the relationship between an observed outcomes and control variables. While not uncontroversial, these assumptions seem reasonable in an analysis intended to examine bias due to sorting.

It is this neighborhood quality index θ_j for which individuals will implicitly be willing to pay higher price of entry p_j to enter a given neighborhood j.[20]

Structure of Equilibrium. Decisions in an EPS-style model are driven by the trade-off between consuming more of the neighborhood (influenced) good y and the price of entry into the neighborhood p. As p increases the individual has less money available for the consumption of all other (non-neighborhood) goods. In equilibrium, the price of entry into each neighborhood adjusts so as to ration the quality of the neighborhood good θ_j available there.

To derive predictions about the structure of the equilibrium, it is helpful to make the following single-crossing properties on preferences:

$$(5) \qquad \frac{d^2 p}{dyde} < 0; \quad \frac{d^2 p}{dyd\mu} < 0;$$

The first of these single-crossing properties implies that as an individual's financial endowment increases, the slope of the indifference curve between the price of entry into the neighborhood p and the consumption of the neighborhood good y decreases, *ceteris paribus*. The second condition implies that the same holds when an individual's preferences for the neighborhood good increases.

Given these single-crossing assumptions, Epple-Platt (1998) demonstrates that a sorting equilibrium exists and can be characterized by two properties that are relevant for our analysis. The first property is actually a pair of stratification conditions: that (i) conditional on tastes, individuals are perfectly stratified across neighborhoods on the basis of their initial financial endowment e and (ii) conditional on initial financial endowment, individuals are perfectly stratified across neighborhoods on the basis of tastes, μ. These stratification properties can be seen in the following graphical depiction of an EPS equilibrium:

[20] In Section 4 below, we consider generalizations of this simple EPS model to cases where individuals value more about neighborhoods than the direct effect of neighborhood on our outcome of interest.

Figure 1: Stratification of Individuals Across Neighborhoods in EPS Equilibrium

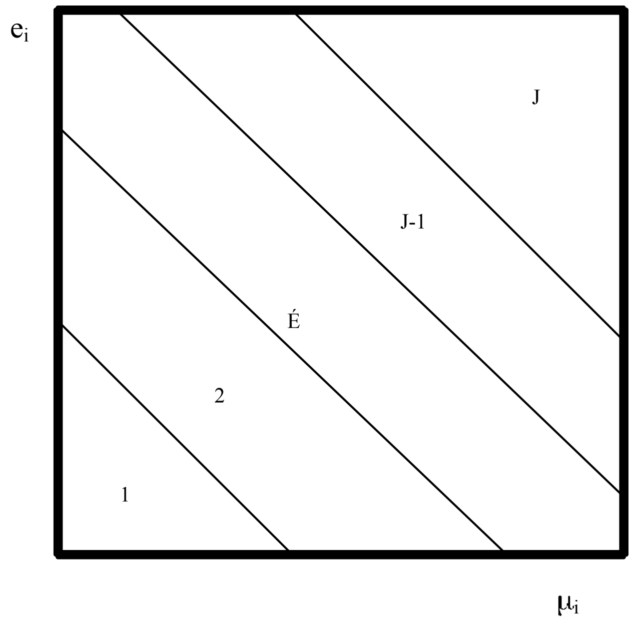

Figure 1 depicts how individuals sort themselves across neighborhoods with increasing values of the neighborhood good $\theta_J > \theta_{J-1} > ... > \theta_2 > \theta_1$. The vertical axis indicates an individual's value of the initial financial endowment *e*, while the horizontal axis depicts an individual's taste for the outcome influenced by neighborhood choice, *y*. The diagonal lines in the figure characterize the boundary in e-μ space that divide the set of individuals that choose one neighborhood versus the other (these boundaries need not be parallel or even straight lines). Finally, for expositional purposes, the graph is drawn assuming a finite support for both tastes and endowments although this is not required.

The stratification result in income can be seen in Figure 1 by considering households with a given value of the taste parameter μ and moving vertically across neighborhoods. In this case, any individual A with endowment greater than individual B $(e_A > e_B)$ chooses a neighborhood with at least as great a value of θ. Likewise, the stratification result in tastes can be seen by conditioning on income and moving horizontally across neighborhoods. In this case, any individual A with tastes greater than individual B ($\mu_A > \mu_B$) chooses a neighborhood with at least as great a value of θ. Notice more generally, that not only are endowments and tastes positively correlated across neighborhoods, but they are negatively correlated within neighborhoods. That is, conditioning on a particular neighborhood the individuals with the highest endowments systematically have the lowest expected level of tastes and vice versa.

Returning to our main estimating equation (1), the key insights that the EPS model provides regarding the resulting correlation between observed and unobserved attributes on the right hand side can be seen if we assume that individual tastes μ are positively correlated with both observed and unobserved productive individual attributes: Z_i, ω_i. Specifically, assuming that:

$$(6) \qquad E[\mu_i'\omega_i] > 0 \quad and \quad \beta_1 E[\mu_i' Z_i] > 0$$

In this case, the structure of the EPS equilibrium implies two key conditions:

$$(7) \qquad \begin{array}{l} (i) \quad E[\mu_i'\theta_j] > 0 \;\Rightarrow\; \beta_1 E[Z_i'\theta_j] > 0 \;\Rightarrow\; \beta_1 E[Z_i'\xi_j] > 0; \\ (ii) \quad E[\mu_i'\theta_j] > 0 \;\Rightarrow\; \beta_2 E[\mu_i' X_j] > 0 \;\Rightarrow\; \beta_2 E[\omega_i' X_j] > 0; \end{array}$$

Condition (7.i) implies that productive observable individual attributes are positively correlated with unobserved neighborhood attributes. Imagine, for example, that an individual's educational attainment both positively affected her taste for the neighborhood good and had a positive direct influence on y. In this case, we would expect a positive correlation between individual educational attainment and any unobserved aspects of a neighborhood that contributed to the production of y. Condition (7.ii) implies that the reverse also holds: that observable neighborhood attributes are correlated with unobserved individual attributes. Returning to our educational attainment example, imagine if educational attainment were unobserved. In this case, we would expect a positive correlation between this unobserved individual attribute and any observed attributes of a neighborhood that contribute to the production of y.

Another property that follows directly from the EPS model is:

$$(8) \qquad E[e_i'\mu_i \mid j] < 0 \;\Rightarrow\; E[e_i'\omega_i \mid j] < 0$$

that *conditional on neighborhood*, endowments and tastes are negatively correlated. From this correlation, it is easy to understand why including neighborhood fixed effects when estimating equation (1) does not provide unbiased estimates of either observed individual attributes of the effect of neighborhood. That is, when the analysis is restricted to within-neighborhood comparisons of outcomes, those individuals within a neighborhood observed with the highest values of observed endowments systematically have lower tastes for the neighborhood good. So, again, to the extent that individual tastes are positively correlated with individual productivity, correlation between the observed and unobserved components of equation (1) would remain.

Broad Estimation Strategy. Our estimation strategy consists of two main parts designed to break the complex correlation patterns between the observed and unobserved components of equation (1). As the correlations in (7) and (8) make clear, the sorting bias arising from the presence of both individual and neighborhood unobservables in equation (1) is substantially more complex that simple selection problems. For example, if an individual with apparently high tastes for neighborhood quality based on their observables resides in a apparently low quality neighborhood based on location observables, the underlying reason for this sorting outcome is unclear – either the neighborhood has unusually high unobservables or the person has unobservables associated with a very low preferences for neighborhood quality.

The key insight that we draw from the EPS sorting model in this paper is that this apparent ambiguity can be resolved by using information on neighborhood prices. If the neighborhood in question has a very high price, the natural conclusion is that neighborhood unobservables are very good, but if the price is low the individual taste unobservables must be negative. In this way, we will use a flexible function of housing prices as a proxy for the unobserved portion of neighborhood quality in (1), which reduces the problem to a more traditional problem arising from selection into neighborhoods based on individual unobservables. We then address this more traditional selection problem using standard instrumental variable techniques to break the correlation between neighborhood attributes and individual unobservables.[21]

Estimation Strategy - Part I. To illustrate this two-part solution, we begin by considering the problem described in condition (7.i) above: the correlation of observable individual attributes with unobserved neighborhood attributes. In this case, it turns out that a second property of the EPS equilibrium suggests a natural fix. In particular, EPS prove that neighborhoods with increasing values of $\theta_J > \theta_{J-1} > ... > \theta_2 > \theta_1$ are also ordered monotonically in terms of neighborhood housing prices: $p_J > p_{J-1} > ... > p_2 > p_1$.

This monotonicity condition implies that there exists a function f such that $\theta = f(p)$. Thus, a non-parametric function of neighborhood housing prices can serve as a perfect control function for θ in equation (1):

[21] This solution requires that the supply elasticity of group openings be constant across groups. In the neighborhood effects model, the supply elasticity of housing must be constant across neighborhoods, which suggests that price might not be a suitable control function for an analysis including exurban and rural areas with high elasticities of supply. In our application, we focus on the Boston Metropolitan Statistical Area, which is heavily developed with little opportunity for the construction of new housing.

(9) $$y_{ij} = \beta_1 Z_i + f(p_j) + \omega_i + \varepsilon_{ij}$$

Given a consistent estimate of *f(p)*, β_2 can then be recovered from a simple regression of *f(p)* on *X*:

(10) $$\hat{f}(p_j) = \beta_2 X_j + \xi_j$$

A difficulty remains, however, if equation (9) were estimated via OLS: namely the correlation of $f(p_j)$ with ω_i. The remaining correlation between the observed and unobserved portion of equation (9) is due directly to the stratification result in the EPS model: simply put, individuals with high tastes for the outcome of interest choose higher quality neighborhoods and, as a result, to the extent that these tastes are correlated with the direct effect of individual attributes bias the estimation of equation (9). But this problem is a standard selection problem that can be addressed by finding an instrument that is correlated with the price of an individual's neighborhood but not with the individual's unobserved attribute. In the next section, we describe our proposed solution to this problem. The key insight to take away from the analysis of the sorting equilibrium is that prices can serve as a control function for the neighborhood component of the unobservable in the outcome equation (1), thereby reducing the identification to the more manageable one of dealing with a single unobservable.

4. Estimation Details

As just described, our broad estimation strategy is divided into two main parts: (i) including a control function based on housing prices in the main estimating equation and (ii) instrumenting for this control functionand other exogenous neighborhood attributes with variables that are correlated with the control function but not the individual unobserved attribute ω. In this section we provide the details of the implementation of these two main parts of the strategy. We then extend the analysis to allow for endogenous neighborhood attributes such as the socioeconomic characteristics of one's neighbors. Finally, we conclude this section by considering generalizations of the simple EPS sorting model presented above, discussing the extent to which the estimation strategy can be extended in those circumstances.

Implementing the Control Function. In practice, we make one key modification when including a control function in equation (9). Specifically, instead of generating a control function for the full neighborhood quality index θ, we instead focus on developing a control function for just the unobservable, leaving $\beta_2 X_j$ in the main estimating equation. Specifically, we estimate a control function for ξ_j as the average residual for each neighborhood arising from a simple housing price equation estimated for the entire metropolitan area. The housing price (p_{ij}) can be described by

$$(11) \qquad p_{ij} = \delta_1 W_{ij} + \delta_2 X_j + \lambda_j + u_{ij}$$

where W_{ijm} is a vector of housing unit attributes. Controlling for housing characteristics absorbs out any aspect of prices that are explained by housing attributes. We do this because we think housing attributes are a dimension of prices that are unlikely to contribute directly to labor market outcomes. We then estimate:

$$(12) \qquad y_{ij} = \beta_1 Z_i + \beta_2 X_j + \beta_4 \hat{\lambda}_j + \omega_i + \varepsilon_{ij}$$

where we must deal directly with the correlation of X_j and $\hat{\lambda}_j$ with ω_i.

Instrumenting for Neighborhood Attributes. To address this correlation we want to instrument for X_j and λ_j with a portion of observed neighborhood unobserved attribute that is uncorrelated with an individual's own unobserved attribute. We propose to use a function of the average values of observed neighborhood prices for families with the same observable characteristics Z_i as instruments: $E[X_j, \hat{\lambda} | Z]$. The logic behind these instruments is that (i) the instruments should be predictive of location because similar individuals should make very similar location choices if they face the same opportunity set (metropolitan housing market), and (ii) the instruments should not be correlated with the individual's unobservable because they are based entirely on individual observables that have already been included directly in the labor market equation.[22]

Most individual and family attributes, such as parent's education, or family size, are discrete in nature. For the purpose of developing these instruments, we organize individuals into homogenous cells based on all possible permutations of the discrete observable attributes that explain an individual's outcomes in the labor market. Specifically, the mean neighborhood

[22] In principle, one might imagine that individuals in the same cell are similar on unobserved features, such as ability or tastes, so that the cell members location choices are driven by unobservables that are similar to the unobservables that drive the individual's location choice. This possibility is ruled out, however, by the assumption in equation (2) that individual observables are uncorrelated with individual unobservables.

exposure within an individual's cell is used to instrument for the individual's actual exposure to various neighborhood attributes. The reader should note that an individual's actual location attributes are always excluded from the calculation of the cell exposure rates applied to a specific individual.

A couple of additional features about this instrument are worth noting. First, notice that $E[X_j, \hat{\lambda} | Z]$ are essentially nonparametric predictions of theobserved and unobserved quality of neighborhood that an individual with a particular set of characteristics Z would choose. In this way, our IV approach amounts to using a fully non-parametric sorting model to predict each individual's neighborhood attributes given her observable characteristics. This empirical strategy exploits the non-linearities inherent in the sorting process. That such non-linearities could serve as the basis for identification of individual outcome equations in the presence of sorting has been key insight of the work by Brock and Durlauf (2001, 2002, 2005) and has been exploited in closely related work by Bayer and Timmins (2006). Ekeland, Heckman, and Nesheim (2004), Bajari and Benkhard (2005) and Bajari and Kahn (2005) use similar sources of identification in the estimation of hedonic models.

Second, notice that in a single metropolitan housing market this expectation relies on non-linearities. If Z were allowed to enter (1) completely flexibly, the instrument would contain no independent variation. In the application that follows, which is based on data from a single large metropolitan area, the independent variation in our instrument derives from the fact that we simultaneously use multiple household characteristics to define the cells upon which are instruments our based. At the same time, we include each type of characteristic (e.g., education, household structure) only directly in the outcome equation (1). The effect of neighborhood would be unidentified if the outcome model included a fixed effect for each of cell of observationally equivalent individuals. It is important to point, however, that the method that we propose here could easily be extended to multiple metropolitan areas. In that case, even if fixed effects were included directly in (1) for each household category upon which the instrument was based, the instrument would have independent variation due to variation in average location decisions made by identical household types in different metropolitan markets.[23]

Allowing for Endogenous Neighborhood Attributes. A final endogeneity issue arises in our application because some of the neighborhood attributes that we would like to consider are endogenously determined by the sorting process itself. Specifically, in our baseline

[23] For an example, see Ross and Zenou (2005).

specifications, we include measures of the average educational attainment and percent of households in poverty within the neighborhood in equation (1). Re-writing equation (11) and (12) here to explicitly include neighborhood averages of certain individual attributes $\overline{Z}_j$ gives:

(13) $$p_{ij} = \delta_1 W_{ij} + \delta_2 X_j + \delta_3 \overline{Z}_j + \lambda_j + u_{ij}$$

(14) $$y_{ij} = \beta_1 Z_i + \beta_2 X_j + \beta_3 \overline{Z}_j + \beta_4 \hat{\lambda}_j + \omega_i + \varepsilon_{ij}$$

Since the sorting process generates a correlation between Z_i and ξ_j it follows immediately that $\overline{Z}_j$ and ξ_j will be correlated in an analogous way in equation (13). To estimate equation (13) therefore, we develop instruments for neighborhood demographic variables using the composition of neighborhoods with similar fixed or exogenous attributes, such as the employment access of the location or the physical quality of the housing stock in the neighborhood. Since neighborhood attributes tend to be continuous variables, a distance measure is developed to characterize the degree of similarity between neighborhoods. The instruments for each $\overline{Z}_{jm}$ are a weighted average of the $\overline{Z}_{km}$'s for neighborhoods that are similar to neighborhood *j* with the weight based on the degree of similarity or proximity (inverse of the distance in attribute space). Specifically,

(15) $$\hat{Z}_j = Mean_{k \in \Pi_{-jm}} (\overline{Z}_k W(X_j, X_k))$$

where *W* represents a weighting function based on a non-parametric kernel smoother, such as the tri-cubic kernel where $W(X_j, X_k) = (1 - (D(X_j, X_k)/D_{Max})^3)^3$, *D* is a distance function, and D_{max} is the maximum distance over which neighborhoods will be considered, see McMillen (1996). The instrument is exogenous to $\overline{Z}_j$ given the exogeneity of X_j.[24] [25]

[24] The cubic spline requires the specification of a maximum distance at which all locations beyond that distance have zero weight. This distance was chosen for each block group so that ten percent of all block groups are used to calculate the average for a given block group. Results are very similar using twenty or five percent of all block groups. Naturally, the block group itself is not included in this weighted average.

[25] Again, as in the use of aggregation to form instruments for the earlier part of our estimation strategy, the use of aggregation in a single metropolitan housing market again implies here that the independent variation in our instrument derives from nonlinearities. It is again important to point, however, that the method that we propose here could easily be extended to multiple metropolitan areas, where again independent variation in the instrument would arise naturally due to across market variation. See the *Identification* sub-section below for more discussion of this point.

Having estimated equation (13), we then estimate equation (14) using the same strategy outlined in the previous sub-section, forming instruments based on average neighborhood attributes for households in the same cell for both exogenous and endogenous neighborhood attributes.

Generalizing Our Simple EPS Sorting Model. A key assumption underlying the simple EPS sorting model that we outlined in Section 3 is that individuals care about their neighborhood choice through only two channels: the neighborhood (influenced) good y and the price of entry into the neighborhood p. It is for this reason that a flexible function of neighborhood housing prices makes a perfect control function for θ, the neighborhood contribution to the production of y. To the extent that households instead value multiple dimensions of neighborhood quality, a flexible function of neighborhood housing prices will no longer serve as a perfect control function for θ.

So, how severe of a problem is this for our proposed methodology? The first thing to note is that if other dimensions of neighborhood quality that affect household consumption are observable, they can be first conditioned out of neighborhood housing prices in a first stage hedonic price regression.[26] This is the reason, for example, that we condition on housing attributes in estimating equation (11) and separately estimate the effect of observed neighborhood attributes on labor market outcomes. If, on the other hand, households value another dimension of neighborhood quality that is unobserved, the control function approach that we propose will no longer work perfectly. In that case, our proposed method will work only as well as housing prices are indeed correlated with that aspect of neighborhood quality that affects the outcomes of interest. In general, we hope that we are able to condition on enough of what might affect housing prices other than neighborhood quality (e.g., housing attributes) directly in the estimation of the first-stage hedonic price regression.

Robustness and Identification. It is important that the reader be aware of the strengths and limitations of this identification strategy. The instruments used for neighborhood contribution in both the labor market outcome and housing price/rent models make intuitive sense. In the individual sample, the exposure of observationally equivalent individuals are used to instrument for the individual's exposure to specific neighborhood attributes, and similarly the demographic

[26] If these other neighborhood attributes are exogenous, this first stage regression can be estimated via OLS. If they are endogenous, instruments would need to be used in the first-stage regression analogous to those discussed in the previous sub-section of the paper.

composition of neighborhoods with observationally similar environmental variables, primarily housing stock composition, is used to instrument for a neighborhood's demographic composition in the sample of housing units. Since these instruments are based on observable characteristics of individuals and neighborhoods, they should be orthogonal to individual and neighborhood unobservables, respectively.

As discussed earlier, the instruments exploit the highly non-linear relationship that is likely to arise between observable attributes and sorting outcomes. The models are identified because some non-linear terms are excluded from the second stage labor market and housing price regressions. We attempt to address concerns with this identification strategy in a number of ways. First, the labor market models are expanded to include important non-linearities, i.e. the interaction of gender with family structure. Further, we rerun the analyses dropping individuals with high levels of human capital with the expectation that these individuals benefit less from neighborhood level information on the labor market. Both of these changes substantially modify the source of identification, and we would expect the results to be unstable and move in unexpected directions in response to these changes. Similarly, we conduct additional analyses that control for the actual neighborhood housing stock composition in the housing price and labor market equations. As above, we would expect spurious estimates to be quite sensitive to including such variables, which are likely to soak up a substantially amount of information associated with neighborhood unobservables.

We also posit that the influence of neighborhood on household capital income is likely to be much smaller than the neighborhood effect on labor market outcomes. We regress outcomes concerning capital income on the same set of individual and neighborhood variables using both ordinary least squares and our instrumental variables specification. If our identification strategy is valid, we would expect that neighborhood variables exhibit a high correlation with capital income using OLS models due to sorting, but much smaller effects using our IV specification.

Finally, the reader should be aware of the implications of the key exogeneity assumptions made in equation (3). The exogeneity assumption for individual variables Z_i is fairly straight forward and well understood in the literature. The impact of an individual variable like education level is likely to capture the influence of both education and any individual unobserved attributes, such as motivation, that are correlated with education. The exogeneity assumption for neighborhood variables is similar for a variable that is considered fixed X_j. For example, good job access may be correlated with some negative aspects of neighborhood quality, and therefore capture both positive effect of job access and the ambiguous effect of the portion of unobserved neighborhood quality that is correlated with job access in the population of neighborhoods.

5. Sample, Control Variables, and Geography

The sample of prime age adults (age 25 to 59) are drawn from confidential Long Form files of the 1990 Decennial Census for the Boston Metropolitan Statistical Area (MSA). The sample drops a small number of non-Hispanic individuals whose race is not defined as white, African-American, or Asian and Pacific Islander, as well as households residing in census tracts where employment access is not defined resulting in a sample of approximately 178,000 individuals.[27]

The bulk of the analysis considers three variables to describe labor market outcomes: labor force participation last week, average number of weeks worked last year conditional on working any weeks, and average hours worked per week last year conditional on having worked at least 40 weeks per year. Three additional labor market variables are also considered that are likely to be behaviorally related to the preceding variables: whether the individual worked any weeks last year, employment last week conditional on being a participant in the labor market, and hours worked last week if employed, see Table 1. It also should be noted that the exact sample for individual outcome variables varies because individuals are dropped from the analysis sample when an outcome is imputed.

For the purpose of describing employment outcomes as well as identifying observationally equivalent individuals, adults in the sample are described by series of categorical control variables (Z) capturing the individual's education (4), age (3), race and ethnicity (4), household structure (6), gender (2), and immigration status (3) where the numbers in parentheses represent the number of categories. The labor market models also contain key interactions of gender with marital status and presence of children to address well-known aspects of female labor force participation in the United States. These variables are also used to create categories based on all permutations of the categorical variables giving rise to 1,718 cells. All prime age adults that belong to the same cell (Ω) as the individual (excepting the individual and their family members of course) are used to calculate average neighborhood attributes. The sample contains households falling into 1,632 cells, and after dropping cells with less than 10 households to reduce measurement error the final sample contains households in 996 cells. This restriction reduces the sample by less than 3,000 individuals and has no effect on any of the empirical results presented in the paper.

[27] The sample contained approximately 700 non-Hispanic individuals who did not fit into one of these racial categories. About 250 individuals resided in block groups where employment access is not defined. See Bayer, Ross, and Topa (2004) for more details on the confidential census data.

Each household and its members reside in a housing unit, and the location of that unit is geo-coded to one of approximately 2,600 census block group in the Boston Metropolitan Area. The neighborhood is described by the following block group characteristics: percent of households in poverty, percent of individuals who are college graduates, percent individuals who are disadvantage minorities (African-American or Hispanic), and percent of individuals who were not born in the U.S.; as well as a job access measure calculated at the census tract level. The job access measure is based on an average of jobs in the same age and education category as the individual where the average is weighted based on the average commute time between the individual's residence and potential employment locations. The weights are based on the coefficient estimates arising from a gravity model, see O'Reagan and Quigley (1998).[28]

A proxy for unobserved neighborhood attributes is calculated as the block group mean residual from a housing price hedonic regression. These residuals are obtained by regressing the logarithm of house price and/or rent (depending upon whether owner-occupied or not) on the physical attributes of each unit: number of bedrooms, number of rooms, age of the unit, whether the unit is single family, whether a multi-family with 10 to 19 units, and whether multi-family with 20 or more units, as well as the neighborhood composition variables described above.[29] As discussed earlier, the neighborhood composition arises from a household sorting process and is endogenous to location unobservables. Therefore, the housing price/rent equation is estimated using instrumental variables, and the instruments are constructed as weighted averages of the demographic composition of similar neighborhoods based on the following neighborhood housing stock variables: percent owner-occupied units, percent single family units, percent large multi-family units (greater than 20 units), percent 1 bedroom or studio units, percent 4 plus bedroom units, average age of housing stock, presence of group quarters, as well as employment access are used as instruments.

---

[28] The gravity model is estimated by regressing the logarithm of the number of workers commuting between two locations on the logarithms of the workers at the origination, of the jobs at the destination, and of the commute time between those locations. Typically, location combinations are dropped when no flows are observed between two locations, which can lead to a noisy measure of employment access at the census tract level. In order to mitigate this noise, we use the logarithm of one plus the flows and impute commute times using a non-parametric kernal smoother based on the cubic spline.

[29] The model allows hedonic attributes to vary by owner-occupancy, and the logarithmic transformation allows the difference between monthly flows (rent) and value (house value) to be captured by the owner-occupancy dummy. A common dummy variable is estimated for each neighborhood using all housing units in that neighborhood whether rental or owner-occupied.

6. Empirical Results

Baseline Models

Prior to discussing and interpreting the empirical results, it is important to acknowledge that a relationship between labor market outcomes and neighborhood attributes may exist for a variety of reasons. The most commonly discussed mechanism involves information barriers to job search and the significance of informal job market referrals. Residential locations that are far from employment concentrations or have high concentrations of individuals who are not strongly attached to the labor market may provide job searchers with little opportunities for mentoring or for gathering information concerning potential job openings. In policy discussions, the usual presumption is that these factors are much more important for youth and low skill workers. On the other hand, a high quality neighborhood may provide the individual with neighborhood amenities that are complementary to leisure or may expose individuals to lower risk of adverse events that influence labor market productivity or behavior. For example, Kling, Liebman, Katz, and Sanbonmatsu (2004) find that public housing residents who were randomly selected to receive vouchers to move to low poverty rate neighborhoods had improved health outcomes. These last two factors may affect all workers equally or may even have a larger effect on high human capital workers with substantial experience or education.

Table 2 presents the results for the OLS and IV estimations of the relationship between individual and neighborhood attributes and being in the labor market, weeks worked last year if working last year, and average hours worked per week if worked at least 40 weeks last year, respectively.[30] The specifications presented control for individual attributes plus employment access, poverty rate, and percent of residents who graduated with a college degree from a four-year institution. The IV specification also includes a control for neighborhood unobservables based on housing prices and rents in each block group.

Focusing on the estimates for neighborhood variables, the estimated impact of neighborhood attributes are substantially larger than the OLS estimates. Specifically, the positive impact of employment access increases dramatically for all three employment outcomes so that a one standard deviation in employment access implies a two percent increase in the likelihood of labor force participation, a one and a third of a week increase in number of weeks worked in a year, and a two and a half hour increase in hours worked per week. The negative impact of a one standard deviation increase in the poverty rate is a seven percent lower labor force participation

[30] The estimates for individual attributes also appear reasonable. Focusing on labor force participation, males have higher participation rates, participation falls between 45 and 59, participation rises with education, increases for married males especially with kids, and decreases for married females especially with young children.

rate and one week less work during the year with the impact on hours being positive and statistically insignificant. The neighborhood unobservables are also associated with more labor force participation, weeks, and hours.

The percent college educated is negatively associated with all three outcomes. This finding is consistent with previous findings that labor market referrals are used less intensively by individuals with higher levels of education (Ionnides and Loury, 2004) and that college educated individuals may both benefit less from and contribute less to informal job networks (Bayer, Ross, and Topa, 2004). Alternatively, percent college educated may capture local amenities that are complementary to leisure and non-market home production activities. For example, individuals residing in locations with neighbors who have a higher level of education may simply enjoy working less and spending more time at home. As discussed, this explanation might help explain why Moving to Opportunity finds a positive impact of neighborhood on health, but no impact of neighborhood on labor market outcomes. Presumably, the lower poverty rates lead to superior health outcomes and an associated increase in labor market potential, but the exposure to more college educated individuals decreases labor market outcomes.

The finding that OLS estimates of neighborhood effects are biased downwards is consistent with the hypothesis that individuals with poor unobservables in terms of labor market outcomes compensate for these unobservables by sorting into locations with better employment prospects. In the neighborhood effects literature, researchers have often expected to find positive selection where high quality workers reside in high quality locations. While this view makes considerable sense when considering the demand for neighborhood amenities related to quality of life, it is less clear that positive selection will arise on variables that impact labor market participation, such as employment access or the quality of informal job networks. High skill workers with strong attachment to the labor market may be less willing than workers with weak labor market attachment to give up neighborhood quality of life amenities in exchange for access to urban environments with good labor market information and low job search costs.

Table 3 presents the results for alternative education subsamples with the first panel presenting the full sample results and the next two panels containing subsamples after dropping individuals with four-year college degrees or dropping individuals with two or more years of college, respectively. The effect of employment access and poverty on labor force participation increases in magnitude as high human capital individuals are eliminated from the sample. This pattern should be expected if the influence of employment access and poverty on labor force participation is driven primarily by neighborhood contributions to job networks. Similarly, the effect of poverty on weeks worked increases in magnitude, and the effect of poverty on hours becomes negative but is still insignificant.[31] On the other hand, the negative effect of employment access on weeks per year and hours per week worked is quite stable as college educated workers are dropped from the sample. This result is not very surprising. The models are estimated for people who are already in the labor market so that the influence of job access is likely to represent costs associated with commuting to an existing job. Commuting costs are often primarily time costs, which actually rise with human capital levels.[32] Table 4 presents a similar exercise dropping white collar workers and shows that the importance of employment access for labor force participation is larger for non-white collar workers.

The negative effect of percent college educated on labor force participation falls for lower skill populations. This effect might be expected to increase in magnitude if this relationship was driven by the availability of job market referrals since non-college graduates would appear to be least likely to benefit from referrals provided by college graduates. The decline in the variable's effect for populations with lower human capital may reflect a lower demand for these neighborhood amenities among low human capital individuals and therefore less substitution towards leisure among non-college educated. Again, Table 4 mirrors the results for education with non-white collar workers experiencing a smaller negative relationship between the presence of college graduates in a neighborhood and labor force participation.

Table 5 presents estimates for subsamples based on gender and family structure. The table focuses on a series of subsamples that are designed to represent increased attachment to the labor market by first dropping married females with children from the sample, then dropping all married females, and finally dropping all females from the sample and focusing only on prime-age males. The results are quite striking. All estimates for the four neighborhood variables decline in magnitude and many become statistically insignificant suggesting that women and especially married women are driving our findings. As in Tables 3 and 4, this table further supports the idea that neighborhoods matter most for the labor market activity of individuals who are not strongly attached to the labor market.

Decomposing the Effects of the Identification Strategy

Table 6 presents the estimates on the neighborhood variables for a series of specifications. The first column presents the results from OLS, and the second column presents

[31] This suggests a larger positive effect of poverty on hours for the college educated. This finding may represent a neighborhood amenities story with high education individuals disliking spending time at home when they reside in high poverty rate neighborhoods and responding to this dislike by working more hours.
[32] See Ross and Zenou (2004) for a study that examines the relationship between commute time and labor market outcomes.

the results from a second stage estimation where the neighborhood fixed effects from labor market models are regressed upon neighborhood variables. The third column contains estimates for a simple instrumental variable model where the three neighborhood variables are predicted using the expected exposure level based on observationally equivalent individuals. The final three columns add a housing price residual from a simple housing price hedonic using ordinary least squares, instrument for that residual based on observationally equivalent individuals, and finally instrument for a unbiased residual arising from using IV in the housing price/rent model.

The main conclusion arising from this table is that the increase in the importance of neighborhood variables arises from instrumenting for those variables in order to break the link between those neighborhood variables and the individual unobservable. The two stage fixed effect estimates look nothing like the results from the IV specification, and the IV specifications are broadly similar in terms of the effect of observed neighborhood attributes. In addition, the housing price residual does not matter until an instrument is used to break the correlation between those neighborhood unobservables and individual unobservables. The overall effect of neighborhood appears to be smaller in the final IV specification as compared the intermediate IV specifications suggesting that the effect of neighborhood may in some cases be overstated when the model does not correctly control for sorting over location specific unobservables.

As discussed earlier, these findings are consistent with a compensation strategy where individuals with lower likelihoods of employment seek out neighborhoods that provide the best opportunity for employment. Of course, the negative correlation between individual labor market unobservables and neighborhood contribution to labor market outcomes may be driven by tastes over neighborhood attributes. For example, individuals with poor labor market unobservables may also exhibit the weakest preference for positive amenities associated with neighborhoods that have poor job access or attract a large number of college graduates based on their housing stock, and as a result these individuals reside in neighborhoods that provide better job market opportunities. On the other hand, the influence of location unobservables appears to arise from positive selection where individuals with high taste observables reside in neighborhoods with positive neighborhood unobservables in terms of labor market outcomes.

Exploring Neighborhood Determinants

Table 7 presents a series of specifications starting with no neighborhood controls except for the housing market residual and then expanding the list of controls to add poverty, employment access, percent with a four-year college education, percent disadvantaged minorities, and finally percent not born in the United States in sequence. A unique set of neighborhood housing price residuals is constructed for each specification where the residual is conditional on the same set of neighborhood controls that were included in the labor market equation. For example, in the no neighborhood control specification, the housing price regression contains no neighborhood controls, and the housing price residual captures the net impact of all aspects of neighborhood quality that are reflected in housing prices.

The key finding of a large negative impact of poverty on labor force participation and weeks worked is quite robust across specifications. The estimated coefficients are similar in magnitude whether or not the specification includes employment access and percent college educated and the magnitude increases with the inclusion of the share minority and immigrant because those neighborhood variables, especially share immigrant, appear to be associated with higher levels of work on all three measures. Neighborhoods with a high share of immigrants may provide especially fertile ground for job referrals and other aspects of the informal job search process. The positive impact of employment access on weeks worked and hours appears robust, but the magnitude falls off as the share minority and immigrant variables are included, and employment access appears to have no impact on labor force participation after including the minority and immigration composition variables.[33] The negative relationship between percent college educated and labor market outcomes is very stable for all three outcome variables.

The estimated coefficient on neighborhood quality is smaller in magnitude for all three labor market outcomes and negative for weeks and hours worked in the model that does not contain any other neighborhood variables. In this model, the neighborhood quality variable captures the net effect of neighborhood given the correlation between different factors that arise in equilibrium, and this net affect appears to be smaller than the individual effects of neighborhood attributes and ambiguous in sign. In equilibrium, the share of college graduates is negatively with poverty rates, and yet both variables reduce the rate and intensity of labor force participation. In practice, they likely cancel out leading to little net influence of neighborhood quality (as captured by price) on labor market outcomes. Once the college degree variable is included, the sign on the housing price residual is consistently positive and the estimated magnitudes are quite stable. Whether the variable captures the low referral contribution of college graduates or consumption amenities that increase the demand for leisure, the inclusion of this variable separates two sets of neighborhood unobservables that are both positively correlated with

[33] The fact that the employment access estimates may not be robust to the inclusion of additional neighborhood variables should not be surprising. Remember, unlike the neighborhood demographic composition variables, job access is assumed exogenous to neighborhood unobservables, which will lead bias due to omitted minority and immigration variables if job access is correlated with omitted neighborhood variables that attract those populations.

price based measures of neighborhood quality, but have conflicting impacts on labor market outcomes.

Additional Validation and Robustness Efforts

Table 8 presents estimates of the relationship between capital income and neighborhood variables in order to see whether our identification strategy implies unrealistically large impacts of neighborhood attributes on capital income. Such findings would suggest that our identification strategy is flawed. Capital income is very noisy and attempts to estimate linear models of capital income did not provide credible estimates on individual attributes. For example, these analyses found no statistically significant relationship between age or education and capital income. In order to mitigate the effect of noise in the self-reported capital income, we focus on three binary variables, which were defined as zero if the individual had between zero capital income and some positive threshold, one if they had capital income above that threshold, and missing if capital income is not reported, imputed or negative. The three thresholds used are \$0, \$1,000, and \$3,000.

Employment access, percent college educated, and in some cases poverty are all correlated with capital income as indicated by the simple OLS regressions. The estimates on neighborhood variables from the instrumental variable specifications are always statistically insignificant and almost always smaller that the estimates arising from OLS. The one exception is the coefficient on poverty in the model for whether capital income is above \$3,000. Even for this estimate, the magnitude of the effect is quite small with a one standard deviation in poverty leading to a less than one percent change in likelihood of having capital income above \$3,000.[34]

Table 9 presents the results for three alternative indicators of labor market outcomes: whether worked any weeks last year, whether employed last week, and number of hours worked last week if employed. These variables parallel the three dependent variables used for most of the analysis with worked last year capturing behaviors related to labor force participation, employed last week capturing the risk of unemployment that might reduce the number of week worked in any year, and hours last week capturing behaviors similar to those captured by average hours worked per week last year. The first panel contains the results for the original three outcome variables and the second contains the results for these three alternative variables. The estimated effects of neighborhood attributes based on the original variables and based on the alternative

[34] A reader might question whether the increasingly positive coefficient on poverty in the IV specification might represent a trend and become large and significant for higher thresholds. We examined models with higher capital income thresholds and did not find any such trend.

dependent variables are quite similar.[35]

Table 10 incorporates a control for the quality of the housing stock in a neighborhood, which is an aggregation of the same housing stock variables used to instrument for neighborhood composition variables. The original IV specification and the specification that includes this control for housing stock are shown side by side. A quick comparison confirms that the magnitudes of all estimated coefficients are quite stable to the inclusion of a housing stock control into both the labor market and housing price/rent models. The reader should note that the model includes the actual housing stock rather than an instrument based on the exposure of observationally equivalent households. The inclusion of the housing stock control is intended to assure that the housing price residual is identified by unexplained variation in housing prices rather than a housing stock exclusion restriction, and the large and significant coefficient estimate on housing stock represent sorting bias rather than any direct effect of neighborhood housing stock on labor market outcomes.[36]

7. Summary and Conclusions

In this paper, we consider the general problem of identifying the effect of individual and group attributes on individual outcomes in a model that allows for both individual and group unobservables. We begin by using a simple version of the canonical vertical model of sorting developed in Epple and Platt (1998) and Epple and Sieg (1999) (EPS) to highlight the complex set of correlations that even a simple model of residential sorting induces between all individual and group attributes. We then offer a non-parametric solution to this identification problem that is grounded in the structure of the sorting equilibrium in the EPS model. In particular, we exploit the monotonic relationship between neighborhood housing prices and neighborhood quality in equilibrium to show that a flexible function of neighborhood housing prices serves as a suitable control function for the neighborhood unobservable in the labor market outcome regression. By including this control function, we eliminate the group unobservable from the regression, thereby reducing the problem to a more standard selection problem with a single individual-level unobservable.

To address this more standard selection problem, we use aggregation to develop suitable instruments for both exogenous and endogenous group attributes. Instrumenting for each

[35] The participation and hours variables are directly comparable to each other in magnitude. The estimates in the employment and weeks worked equations are not, but one can verify that the relative magnitudes of the coefficient estimates from the two models are quite close.
[36] Results are also robust to a model that instruments for housing stock, and in that model housing stock is not statistically significant.

individual's observed neighborhood attributes with the average neighborhood attributes of a set of observationally identical individuals eliminates the portion of the variation in neighborhood attributes due to sorting on unobserved individual attributes.

To illustrate our proposed methodology, we estimate a wide variety of labor market models using confidential data on the Boston Metropolitan Area from the 1990 census long form. We find that neighborhood has large and complex affects on labor market outcomes. Employment access, low levels of poverty, a low fraction of college graduates, and high levels of unobserved neighborhood attributes are all associated with higher levels of labor force participation, greater number of weeks worked in a year, and with the exception of poverty greater average number of hours worked per week. The estimated effects of neighborhood variables are economically meaningful with for example a one standard deviation increase in employment access leading to approximately a four percentage point increase in labor force participation in the subsample of individuals who have never attended college. Moreover, the estimated effects are substantially larger than estimates arising from ordinary least squares suggesting that individuals with a lower likelihood of obtaining employment have sorted into locations with superior labor market opportunities potentially to compensate for their poor unobservables. It is notable that the core results in the paper are robust across many outcomes variables and a wide variety of specifications.

As expected, the positive impact of low neighborhood poverty rates and good job access on labor force participation increases as high human capital individuals or white collar workers are deleted from the sample. The existing literature suggests that these individuals are least likely to benefit from informal labor market referral networks. On the other hand, the positive impact of good job access on the intensity of labor force participation as captured by weeks per year and hours per week does not change as the human capital level of the sample falls. This finding may in part be due to the fact that high human capital individuals have a high cost of time and therefore may substitute away from work as commutes increase. The effects over gender are even more striking all findings decline in magnitude and many become statistically insignificant as married women and eventually all women are deleted from the sample. Overall, the results indicate that neighborhood effects are most important for individuals with weak attachment to the labor market, especially married women.

While the effect of individual variables appears large, the net effect of neighborhood quality is actually quite small. This finding appears to be driven by the strong negative effect of the percent of college graduates in a neighborhood on labor market outcomes. Neighborhoods with low poverty rates and other attributes that positively impact labor market outcomes appear correlated with the percent of college graduates in equilibrium. These competing effects lead to small and sometimes negative relationships between overall neighborhood quality and various labor market outcomes, which is consistent with findings in the Moving to Opportunity program that improvements in neighborhoods quality had little or no impact on earnings. Moreover, these findings help explain a puzzle in the MTO results. Voucher recipients in MTO experience improved health outcomes, but do not experience the improvement in labor market outcomes often associated with improvements in physical and mental health. The positive effects of improved health on labor market potential may have been counteracted by other influences of neighborhood that lead to reduced labor supply.

These findings suggest that a richer understanding of the relationship between neighborhood and economic self-sufficiency is required to address the high unemployment rates and low incomes occurring in poor, central city neighborhoods. High poverty rate neighborhoods appear to have a large negative affect on labor market outcomes, especially for low human capital populations. This large effect might be attributable in part to the negative impacts of high poverty locations on health and emotional well being found in the Moving to Opportunity program. Future mobility programs should take into account the possibility that small net effects of neighborhood quality hide large positive and negative impacts on labor market outcomes. For example, the potential negative impact of moving on informal referral networks may in part be offset by increased provision of formal job search support.

References

Addison, John T. and Pedro Portugal. 2002. Job Search Methods and Outcomes. *Oxford Economics Papers*, 54 (3), 505-33.

Aizer, Anna and Janet Currie. 2004. Networks or Neighborhoods? Correlations in the Use of Publicly-Funded Maternity Care in California. *Journal of Public Economics*, 88 (12), 2573-85.

Arcidiacono, Peter and Jacob Vigdor. 2004. On the Value of Inter-Racial Contact and Affirmative Action as a Means to Achieve It. Duke University unpublished manuscript.

Bajari, Patrick, and Lanier Benkard. 2005. Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach *Journal of Political Economy*, 113(6), 1239-76.

Bajari, Patrick, and Matthew Kahn. 2005. Estimating Housing Demand with and Application to Explaining Racial Segregation in Cities. *Journal of Business and Economic Statistics*, 23(1), 20-33.

Bayer, Patrick, Robert McMillan, Kim Rueben. 2005. An Equilibrium Model of Sorting in an Urban Housing Markets. NBER Working Paper No. 10865

Bayer, Patrick, Stephen L. Ross, Giorgio Topa. 2005. Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. NBER Working Paper No. 11019.

Bayer, Patrick and Christopher Timmons. 2005. On the Equilibrium Properties of Locational Sorting Models. *Journal of Urban Economics*, 57(3), 462-77.

Bayer, Patrick and Christopher Timmons. 2006. Estimating Equilibrium Models of Sorting across Locations. *Economic Journal*, forthcoming.

Benabou, Roland. 1993. The Workings of a City: Location, Education, and Production. *Quarterly Journal of Economics*, 108(3), 619-52.

Benabou, Roland. 1996. Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance, *American Economic Review*, 86(3), 584-609.

Bertrand, Marianne, Erzo Luttmer, and Sendhil Mullainathan. 2000. Network Effects and Welfare Cultures. *Quarterly Journal of Economics*, 115 (3), 1019-55.

Blau, David and Phillip Robbins. 1990. Job Search Outcomes for the Employed and Unemployed. *Journal of Political Economy*, 98 (3), 637-55

Blau, David. 1992. An Empirical Analysis of Employed and Unemployed Job Search Behavior, *Industrial and Labor Relations Review*, 45, 738-52.

Bortnick, Steven and Michelle Ports. 1992. Job Search Methods and Results: Tracking the Unemployed, *Monthly Labor Review*, 115(12), 29-35.

Brock, William A. and Stephen N. Durlauf. 2006. Identification of Binary Choice Models with Social Interactions. *Journal of Econometrics*, forthcoming.

Brock, William A. and Stephen N. Durlauf. 2005. A Multinomial Choice Model with Social Interactions. In L. Blume, S. Durlauf eds., <u>The Economy as and Evolving Complex System III</u>, Oxford: Oxford University Press: New York, 175-206

Brock, William A. and Stephen N. Durlauf. 2002. A Multinomial Choice Model of Neighborhood Effects. *American Economic Review*, 92, 298-303.

Brock, William A. and Steven N. Durlauf. 2001. Discrete choice with social interactions. *Review of Economic Studies*, 68, 235-260.

Calvo-Armengol, Antoni and Matthew O. Jackson. 2004. The Effects of Social Networks on Employment and Inequality, *American Economic Review*, 94 (3), 426-54.

Card, David and Jesse Rothstein. 2005. Racial Segregation and the Black-White Test Score Gap, Unpublished manuscript.

Case, Anne C. and Lawrence F. Katz. 1991. The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths. NBER Working Paper No. 3705.

Case, Anne C., Darren Lubotsky, Chris Paxson. 2002. Economic Status and Health in Childhood: The Origin of the Gradient. *American Economic Review*, 92, 1308-1334.

Corcoran, Mary, Linda Datcher and Greg Duncan. 1980. Information and Influence Networks in Labor Markets, in *Five Thousand American Families: Patterns of Economic Progress*, edited by Greg Duncan and James Morgan, 7, 1-37, Ann Arbor, MI: Institute For Social Research.

Cutler, David and Edward Glaeser. 1997. Are Ghettos Good or Bad? *Quarterly Journal of Economics*, 112, 827-872.

Devine, Theresa and Nicholas Kiefer. 1991. *Empirical Labor Economics: The Search Approach*, New York: Oxford University Press.

Ekeland, Ivar, James Heckman, and Lars Nesheim. 2004. Identification and Estimation of Hedonic Models. *Journal of Political Economy*, 112(1.2), S60-S108.

Elliot, James. 1999. Social Isolation and Labor Market Isolation: Network and Neighborhood Effects on Less-Educated Workers. *Sociological Quarterly,* 40 (2), 199-16.

Epple, Dennis. 1987. Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products. *Journal of Political Economy*, 95(1): 59-80.

Epple, D., R. Filimon, and T. Romer. 1984. Equilibrium Among Local Jurisdictions: Towards an Integrated Approach of Voting and Residential Choice*. Journal of Public Economics*, 24, 281-304.

Epple, D., R. Filimon, and T. Romer. 1993. Existence of Voting and Housing Equilibrium in a System of Communities with Property Taxes*. Regional Science and Urban Economics*, 23, 585-610.

Epple, Dennis and Glen J. Platt. 1998. Equilibrium and Local Redistribution in an Urban Economy when Households Differ in Both Preferences and Incomes. *Journal of Urban Economics*, 43, 23-51.

Epple, Dennis and Richard Romano. 1998. Competition between Private and Public Schools, Vouchers, and Peer-Group Effects. *American Economic Review*, 88(1), 33-62.

Epple, Dennis and Thomas Romer. 1991. Mobility and Redistribution, *Journal of Political Economy*, 99(4), 828-58.

Epple, Dennis, Thomas Romer and Holger Sieg. 2001. Interjurisdictional Sorting and Majority Rule: An Empirical Analysis. *Econometrica*, 69(6), 1437-55.

Epple, Dennis and Holger Sieg. 1999. Estimating Equilibrium Models of Local Jurisdictions. *Journal of Political Economy*, 107, 645-681.

Ellen, Ingrid and Margery A. Turner. 1997. Does Neighborhood Matter? Assessing Recent Evidence. *Housing Policy Debate*, 8, 833-866.

Ettner, Susan L., Richard G. Frank, and Ronald C. Kessler. 1997. The Impact of Psychiatric Disorders on Labor Market Outcomes. *Industrial and Labor Relations Review*, 51, 64-81.

Evans, Willian, Wallace Oates, and Robert Schwab. 1992. Measuring Peer Group Effects: A Study of Teenage Behavior. *Journal of Political Economy*, 10, 966-991.

Fernandez, Raquel and Richard Rogerson. 1996. Income Distribution, Communities, and the Quality of Public Education. *Quarterly Journal of Economics*, 111(1), 135-164.

Fernandez, Raquel and Richard Rogerson. 1998. Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education Finance Reform. *American Economic Review*, 88(4), 813-33.

Ferreira, Fernando. 2003. You Can Take It with You: Transferability of Proposition 13 Tax benefits, Residential Mobility, and Willingness to Pay for Housing Amenities. University of Pennsylvania unpublished manuscript.

Ferreyra, Maria. 2003. Estimating the Effects of Private School Vouchers in Multi-District Economies. Carnegie-Mellon University unpublished manuscript.

Granovetter, Mark S. 1995. *Getting a Job: A Study of Contacts and Careers*. Cambridge, MA: Harvard University Press.
Holzer, Harry. 1987. Informal Job Search and Black Youth Unemployment. *American Economic Review*, 77(3), 446-52.
Holzer, Harry. 1988. Search Method Use by Unemployed Youth. *Journal of Labor Economics*, 6(1), 1-20.
Ihlanfeldt, Keith. 2000. The Spatial Mismatch Hypothesis: A Review of Recent Studies and Their Implications for Welfare Reform. *Housing Policy Debate*, 9, 849-892.
Ihlanfeldt, Keith. 1992. *Job Accessibility and the Employment and School Enrollment of Teenagers*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.
Ihlanfeldt, Keith and David Sjoquist. 1990. Job Accessibility and Racial Differences in Youth Employment Rates. *American Economic Review*, 80, 267-276.
Ioannides, Yannis M. and Linda Datcher Loury. 2004. Job Information Networks, Neighborhood Effects, and Inequality. *Journal of Economic Literature*, 42, 1056-93.
Ioannides, Yannis M. and Jeffrey Zabel. 2004. Interactions, Neighborhood Selection and Housing Demand. Tufts University unpublished manuscript.
Jacob, Brian. 2004. Public Housing, Housing Vouchers and Student Achievement: Evidence from Public Housing Demolitions in Chicago, *American Economic Review,* 94(1), 233-58.
Kain, John. 1968. Housing Segregation, Negro Employment, and Metropolitan Decentralization. *Quarterly Journal of Economics*, 82, 175-197.
Katz, Lawrence F., Jeffrey Kling and Jeffrey Liebman. 2001. Moving To Opportunity in Boston: Early Results of a Randomized Mobility Experiment. *Quarterly Journal of Economics*, 116, 607-54.
Kling, Jeffrey, Jeffrey Liebman, Lawrence F. Katz, and Lisa Sanbonmatsu. 2004. Moving To Opportunity and Tranquility: Neighborhood Effects on Adult Economic Self-sufficiency and Health from a Randomized Housing Voucher Experiment. Unpublished manuscript.
Kling, Jeffrey, Jens Ludwig, and Lawrence F. Katz. 2005. Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment. *Quarterly Journal of Economics*, 120, 87-130.
Korenman, Sanders and Susan Turner. 1996. Employment Contacts and Minority-White Wage Differences. *Industrial Relations*, 35 (1), 106-22.
Loury, Linda Datcher. 2003. Some Job Contacts are More Equal Than Others: Earnings and Job Information Networks. Tufts University, unpublished manuscript.
Ludwig, Jens, Greg Duncan, and Joshua Pinkston. 2000. Neighborhood effects on economic self-sufficiency: Evidence from a randomized experiment. Georgetown University unpublished manuscript.
Manski, Charles F. 1993. Identification of Endogenous Social Effects: the Reflection Problem. *Review of Economic Studies*, 60, 531-542.
Marmaros, David and Bruce Sacerdote 2002. Peer and Social Networks in Job Search, *European Economic Review*, 46, 870-879.
Mayer, Christopher. 1996. Does Location Matter? *New England Economic Review*, May/June, 26-40.
Moffitt, Robert. 2001. Policy Interventions, Low-Level Equilibria and Social Interactions. In *Social Dynamics*, edited by Steven N. Durlauf and H. Peyton Young. Cambridge, Mass.: MIT Press.
Montgomery, James D. 1991. Social Networks and Labor-Market Outcomes: Toward an Economic Analysis. *The American Economic Review*, 81(5), 1408-18.
Nechyba, Thomas. 1999. School Finance Induced Migration and Stratification Patterns: the Impact of Private School Vouchers. *Journal of Public Economic Theory*, 1.
Nechyba, Thomas. 2000. Mobility, Targeting, and Private School Vouchers. *American Economic Review*, 90(1): 130-46.

O'Regan, Kathy and John Quigley. 1998. Teenage Employment and the Spatial Isolation of Minority and Poverty Households. *Urban Studies*, 35, 1187-1205

Oreopoulos, Philip. 2003. The Long-Run Consequences of Living in a Poor Neighborhood. *Quarterly Journal of Economics,* 118(4), 1533-75.

Petrin, Amil and Kenneth Train. 2002. Omitted Product Attributes in Discrete Choice Models. University of California-Berkeley unpublished manuscript.

Popkin, Susan J., James E. Rosenbaum and Patricia M. Meaden. 1993. Labor Market Experiences of Low-Income Black Women in Middle-Class Suburbs: Evidence from a Survey of Gautreaux Program Participants. *Journal of Policy Analysis and Management*, 12, 556-73.

Raphael, Steven. 1998. The spatial mismatch hypothesis and black youth joblessness. *Journal of Urban Economics 43, pp. 79-111.*

Rees, Albert and George P. Schultz. 1970. *Workers and Wages in an Urban Labor Market*, Chicago: Univ. of Chicago Press.

Ross, Stephen L. 1998. Racial Differences in Residential and Job Mobility. *Journal of Urban Economics*, 43, 112-36.

Ross, Stephen L. and John Yinger. 1999. Urban Public Finance. In *The Handbook of Regional and Urban Economics, Volume 3: Applied Urban Economics*, edited by Paul Cheshire and Edwin S. Mills. Elsevier Science/North Holland.

Ross, Stephen L. and Yves Zenou. 2004. Shirking, Commuting and Labor Market Outcomes. University of Connecticut unpublished manuscript.

Sacerdote, Bruce. 2001, Peer Effects with Random Assignment: Results for Dartmouth Roommates. *Quarterly Journal of Economics*, 116 (2), 681-704.

Sieg, Holger, V. Kerry Smith, H. Spencer Banzaf and Randall Walsh. 2004. Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods. *International Economic Review*. 45(4), 1047-77

Smith, James P. 1999. Healthy Bodies and Thick Wallets. *Journal of EconomicsPerspectives*, 13(2), 145-166.

Smith, James P. 2003. Consequences and Predictors of New Health Events. NBER Working Paper No. 10063.

Wahba, Jackline and Yves Zenou. 2003. Density, Social Networks and Job Search Methods: Theory and Application to Egypt, CEPR Working Paper #3967.

Walsh, Randall. 2004. Endogenous Open Space Amenities in a Locational Equilibrium. CEA Working Paper No. 04-03, February.

Weinberg, Bruce. 2004. Testing the Spatial Mismatch Hypothesis Using Inter-city Variations in Industrial Composition. *Regional Science and Urban Economics*, 34(5): 505-32.

Weinberg, Bruce. 2005. Social Interactions and Endogenous Association. Ohio State University, unpublished manuscript.

Weinberg, Bruce. 2000. Black Residential Centralization and the Spatial Mismatch Hypothesis. *Journal of Urban Economics*, 48, 110-34.

Weinberg, Bruce, Patricia Reagan and Jeffrey Yankow. Forthcoming. Do Neighborhoods Affect Hours Worked: Evidence from Longitudinal Data. *Journal of Labor Economics*.

Zax, Jeffrey and John Kain. 1991. Commutes, Quits, and Moves. *Journal of Urban Economics*, 29, 153-65.

Zenou, Yves. 2000. Unemployment in Cities. In *Economics of Cities: Theoretical Perspectives*, edited by J. Huriot and J. Thisse. Cambridge: Cambridge University Press.

Table 1: Variable Names, Description, Means, and Standard Errors		
Variable Names	Variable Description	Means and Standard Error
Respondent Outcome Variables		
Labor-Force Participant	One if respondent was working or looking for work at the time of the Census Survey	0.854 (0.352)
Weeks-Worked Last-Year	Total number of weeks worked last year; missing if no weeks worked last year	40.105 19.228)
Weekly-Hours Last-Year	Average number of hours worked per week last year; missing if worked less than 40 weeks last year	34.388 (17.635)
Worked-Last Year	One if respondent worked any weeks last year	0.856 (0.350)
Employed-Last Week	One if respondent was employed last week, zero if unemployed and a labor force participant, and missing otherwise	0.811 (0.390)
Hours-Worked Last-Week	Number of hours worked last week; missing if not employed last week	32.551 19.750)
Positive-Capital Income	One if respondent has positive capital income; missing if capital income negative, imputed or not reported	0.453 (0.498)
Capital-Income >1000	One if respondent has positive capital income; missing if capital income negative, imputed or not reported	0.307 (0.461)
Capital-Income >3000	One if respondent has positive capital income; missing if capital income negative, imputed or not reported	0.268 (0.443)
Categorical Respondent Control Variables		
Male	One if respondent male (omitted category female)	0.482 (0.499)
Age35-44	One if respondent between 35 and 44 years of age (omitted category 25 to 34 years)	0.317 (0.465)
Age45-59	One if respondent between 45 and 59 years of age	0.302 (0.459)
Black	One if respondent non-Hispanic Black (omitted category non-Hispanic white)	0.047 (0.210)
Hispanic	One if respondent Hispanic	0.033 (0.178)
Asian	One if respondent Asian or Pacific Islander	0.029 (0.167)
No-High-School	One if respondent did not graduate from high school (omitted category high school graduate)	0.099 (0.298)
Some-college	One if respondent finished at least two years of college but does not have four year degree	0.247 (0.431)
College	One if respondent gradudated with a four year college degree	0.404 (0.490)
Single-Independent	One if respondent is single and not living with family members (omitted category married not residing with any of their own children who are under the age of 18 – minors)	0.224 (0.417)
Single-parent	One if respondent is a single parent residing with their minor child	0.094 (0.291)
Single-with-family	One if respondent is single and living with family members other than their children	0.051 (0.219)
Married-with-17yr-kid	One if respondent is married and residing with their minor children	0.179 (0.383)
Married-with-0-5yr-kid	One if respondent married residing with their own child under the age of six (omitted category married residing with their own children, but no child under the age of six)	0.187 (0.390)
Married-Female	Interaction between marital status and respondent female	0.321 (0.467)
Married-Female-kids	Interaction between marital status, respondent female, and residing with own minor children	0.184 (0.387)
Married-Female-0-5-kids	Interaction between marital status, respondent female, and residing with own child who is under the age of six	0.093 (0.291)
Non-US-born	One if respondent is U.S. citizen not born in the U.S. (omitted category born in the U.S.)	0.065 (0.247)
Non-US-Citizen	One if respondent is not a U.S. citizen	0.072 (0.259)

Table 1: Variable Names, Description, Means, and Standard Errors (Continued)		
Variable Names	Variable Description	Means and Standard Error
Neighborhood Level Variables		
Employment-Access	Employment access index based on gravity model using non-parametrically smoothed estimates of commuting time between census tracts	1.051 (0.067)
Percent-Poverty	Percent of households in poverty within a census block group	0.051 (0.065)
Percent-College Graduate	Percent of prime age individuals (age 25-59) with a four year college degree within a census block group	0.401 (0.210)
Housing-Price Residual	Block group mean of the housing price residual arising from a metropolitan wide housing price hedonic that controls for the three block group attributes listed above	0.005 (0.067)
Percent-Minority Disadvantage	Percent of households in census block group headed by either an African-American or Hispanic respondent	0.084 (0.177)
Percent-Not-Born US	Percent of prime age individuals in census block group who were not born in the United States	0.069 (0.061)
Housing-Stock Index	Block group mean of a housing stock index based on mean housing stock attributes of each block group using the coefficient estimates on those mean attributes in a housing price hedonic	0.215 (0.167)

Table 2: Models of Labor Models Outcomes						
Variables	Labor Force Participant		Weeks Worked Last Year		Weekly Hours Last Year	
	OLS	Final IV Model	OLS	Final IV Model	OLS	Final IV Model
Male	0.022 (8.12)	0.018 (6.60)	-0.549 (-5.39)	-0.573 (-5.43)	2.664 (28.21)	2.626 (27.10)
Age35-44	0.001 (0.36)	0.009 (3.59)	0.721 (8.77)	1.014 (10.95)	-0.534 (-6.81)	0.050 (0.58)
Age45-59	-0.053 (-22.05)	-0.040 (-11.77)	0.571 (6.51)	1.096 (9.06)	-1.569 (-19.03)	-0.578 (-5.63)
Black	0.031 (5.90)	0.098 (10.70)	0.352 (1.94)	-0.018 (-0.06)	0.707 (4.45)	-1.266 (-4.68)
Hispanic	-0.010 (-1.43)	0.034 (3.74)	-0.365 (-1.48)	-0.670 (-2.12)	0.151 (0.65)	-1.423 (-4.83)
Asian	-0.041 (-4.55)	-0.028 (-3.05)	-0.042 (-0.17)	-0.093 (-0.35)	0.611 (2.31)	-0.074 (-0.28)
No-High-School	-0.087 (-20.44)	-0.076 (-16.14)	-1.271 (-8.52)	-1.535 (-9.02)	0.010 (0.08)	-0.799 (-5.85)
Some-college	0.052 (20.69)	0.054 (17.43)	0.580 (6.58)	1.023 (9.06)	0.411 (5.04)	1.262 (12.31)
College	0.082 (33.70)	0.097 (14.28)	0.672 (7.90)	1.989 (8.12)	1.874 (23.01	4.025 (17.89)
Single-Independent	-0.035 (-12.34)	-0.039 (-8.74)	-1.253 (-11.43)	-1.516 (-9.20)	-1.627 (-13.96)	-2.611 (-16.49)
Single-parent	-0.060 (-16.56)	-0.049 (-11.10)	-2.563 (-17.26)	-2.565 (-14.50)	-3.595 (-28.23)	-3.751 (-25.58)
Single-with-family	-0.148 (-25.83)	-0.126 (-20.97)	-3.443 (-17.66)	-3.446 (-16.63)	-4.082 (-22.54)	-4.517 (-23.50)
Married-with-17yr-kid	0.015 (6.34)	0.023 (8.09)	0.581 (6.14)	0.788 (7.49)	1.068 (9.70)	1.564 (13.18)
Married-with-0-5yr-kid	-0.003 (-1.15)	0.006 (2.36)	0.771 (7.89)	0.945 (8.89)	1.072 (9.57)	1.465 (12.17)
Married-Female	-0.117 (-27.77)	-0.118 (-27.69)	-2.650 (-17.99)	-2.543 (-17.01	-4.506 (-30.99)	-4.363 (-29.47)
Married-Female-kids	-0.059 (-12.20)	-0.059 (-12.10)	-3.565 (-20.65)	-3.518 (-20.34)	-6.551 (-36.33)	-6.485 (-3.00)
Married-Female-0-5-kids	-0.162 (-29.24)	-0.162 (-29.11)	-3.108 (-14.49)	-3.167 (-14.72)	-2.604 (-12.25)	-2.696 (-12.68)
Non-US-born	0.017 (4.12)	0.013 (2.90)	0.513 (3.63)	0.40 (2.53)	0.971 (6.83)	0.583 (3.77)
Non-US-Citizen	-0.020 (-3.91)	-0.034 (-6.00)	-1.857 (-10.66)	-2.082 (-10.15)	0.309 (1.95)	-0.472 (-2.59)
Employment-Access	0.006 (0.32)	0.308 (2.02)	1.769 (2.83)	19.156 (3.41)	-1.501 (-2.68)	36.068 (7.42)
Percent-Poverty	-0.528 (-20.62)	-1.055 (-8.29)	-11.176 (-15.03)	-14.709 (-3.18)	-3.550 (-5.14)	6.027 (1.52)
Percent-College Graduate	-0.099 (-15.01	-0.434 (-9.71)	-2.397 (-12.26)	-13.929 (-8.62)	1.264 (6.11)	-17.186 (-11.93)
Housing-Price Residual		0.514 (9.48)		7.270 (3.76)		11.685 (7.33)

Table 3: Labor Market Models for Subsample with Lower Education Levels						
	Labor Force Participant		Weeks Worked Last Year		Weekly Hours Last Year	
	OLS	Final IV Model	OLS	Final IV Model	OLS	Final IV Model
Full Sample						
Employment-Access	0.006 (0.32)	0.308 (2.02)	1.769 (2.83)	19.156 (3.41)	-1.501 (-2.68)	36.068 (7.42)
Percent-Poverty	-0.528 (-20.62)	-1.055 (-8.29)	-11.176 (-15.03)	-14.709 (-3.18)	-3.550 (-5.14)	6.027 (1.52)
Percent-College Graduate	-0.099 (-15.01	-0.434 (-9.71)	-2.397 (-12.26)	-13.929 (-8.62)	1.264 (6.11)	-17.186 (-11.93)
Housing-Price Residual		0.514 (9.48)		7.270 (3.76)		11.685 (7.33)
Sample After Dropping All Respondents with Degrees from Four Year Colleges						
Employment-Access	-0.004 (-0.19)	0.424 (2.08)	2.616 (3.82)	17.239 (2.37)	-3.901 (-6.00)	44.459 (7.42)
Percent-Poverty	-0.578 (-19.75)	-1.203 (-7.66)	-10.010 (-11.52)	-11.284 (-2.01	-3.020 (-4.01	-0.160 (-0.03)
Percent-College Graduate	-0.092 (-10.37)	-0.225 (-2.94)	-1.671 (-6.26)	-14.480 (-5.39)	-0.403 (-1.56)	-16.394 (-7.39)
Housing-Price Residual		0.522 (7.78)		11.163 (4.58)		11.468 (5.90)
Sample After Dropping All Respondents with Two or More Years of College						
Employment-Access	-0.044 (-1.59)	0.704 (2.28)	3.457 (3.82)	14.367 (1.30)	-4.701 (-5.95)	32.041 (3.59)
Percent-Poverty	-0.604 (-17.59)	-1.533 (-7.00)	-9.034 (-7.97)	-20.073 (-2.53)	-2.104 (-2.41)	-7.825 (-1.19)
Percent-College Graduate	-0.063 (-4.96)	-0.188 (-1.38)	-0.635 (-1.65)	-16.547 (-3.47)	0.156 (0.44)	-8.290 (-2.10)
Housing-Price Residual		0.480 (5.57)		12.046 (3.74)		7.121 (2.88)

Table 4: Labor Market Models for Non-White Collar Subsample						
	Labor Force Participant		Weeks Worked Last Year		Weekly Hours Last Year	
	OLS	Final IV Model	OLS	Final IV Model	OLS	Final IV Model
Full Sample						
Employment-Access	0.006 (0.32)	0.308 (2.02)	1.769 (2.83)	19.156 (3.41)	-1.501 (-2.68)	36.068 (7.42)
Percent-Poverty	-0.528 (-20.62)	-1.055 (-8.29)	-11.176 (-15.03)	-14.709 (-3.18)	-3.550 (-5.14)	6.027 (1.52)
Percent-College Graduate	-0.099 (-15.01	-0.434 (-9.71)	-2.397 (-12.26)	-13.929 (-8.62)	1.264 (6.11)	-17.186 (-11.93)
Housing-Price Residual		0.514 (9.48)		7.270 (3.76)		11.685 (7.33)
Sample After Dropping All White Collar Employees						
Employment-Access	0.033 (2.01	0.599 (3.92)	1.833 (2.59)	19.695 (2.85)	-3.140 (-4.76)	40.048 (7.05)
Percent-Poverty	-0.304 (-14.16)	-0.715 (-5.71)	-11.077 (-12.42)	-11.635 (-2.07)	-3.592 (-4.58)	4.826 (1.05)
Percent-College Graduate	-0.061 (-9.22)	-0.225 (-4.41)	-2.349 (-9.04)	-13.342 (-6.12)	-0.108 (-0.43)	-15.812 (-8.57)
Housing-Price Residual		0.288 (5.56)		10.563 (4.56)		10.586 (5.79)

Table 5: Labor Market Models with Gender and Family Structure Subamples						
	Labor Force Participant		Weeks Worked Last Year		Weekly Hours Last Year	
	OLS	Final IV Model	OLS	Final IV Model	OLS	Final IV Model
Full Sample						
Employment-Access	0.006 (0.32)	0.308 (2.02)	1.769 (2.83)	19.156 (3.41)	-1.501 (-2.68)	36.068 (7.42)
Percent-Poverty	-0.528 (-20.62)	-1.055 (-8.29)	-11.176 (-15.03)	-14.709 (-3.18)	-3.550 (-5.14)	6.027 (1.52)
Percent-College Graduate	-0.099 (-15.01	-0.434 (-9.71)	-2.397 (-12.26)	-13.929 (-8.62)	1.264 (6.11)	-17.186 (-11.93)
Housing-Price Residual		0.514 (9.48)		7.270 (3.76)		11.685 (7.33)
Sample After Dropping All Married Women with Children						
Employment-Access	-0.023 (-1.23)	0.300 (1.88)	0.560 (0.83)	10.518 (1.82)	-2.464 (-4.36)	14.477 (2.82)
Percent-Poverty	-0.548 (-21.49)	-1.283 (-9.60)	-11.728 (-14.178)	-19.112 (-3.97)	-4.700 (-6.48)	3.751 (0.89)
Percent-College Graduate	-0.075 (-11.61)	-0.582 (-12.45)	-2.180 (-10.72)	-14.357 (-8.49)	1.747 (8.29)	-13.214 (-8.66)
Housing-Price Residual		0.468 (8.39)		4.176 (2.12)		4.029 (2.43)
Sample After Dropping All Married Women						
Employment-Access	-0.029 (-0.38)	0.180 (1.09)	0.155 (0.21)	0.363 (0.06)	-3.068 (-5.00)	-4.176 (-0.78)
Percent-Poverty	-0.577 (-21.93)	-1.349 (9.64)	-12.153 (-14.90)	-10.489 (-2.07)	-5.263 (-6.72)	17.288 (3.94)
Percent-College Graduate	-0.049 (-7.65)	-0.273 (5.63)	-1.853 (-8.80)	-5.916 (-3.28)	2.287 (9.95)	1.394 (0.84)
Housing-Price Residual		0.473 (7.98)		4.699 (2.27)		4.680 (2.71)
Sample After Dropping All Women						
Employment-Access	-0.053 (-2.83)	-0.218 (-1.26)	-0.812 (-1.12)	-4.658 (-0.68)	-4.036 (-5.82)	-11.178 (-1.68)
Percent-Poverty	-0.429 (-14.21)	-0.290 (-2.01	-12.042 (-12.45)	-5.867 (-1.01	-4.966 (-4.84)	20.916 (3.94)
Percent-College Graduate	-0.037 (-5.17)	-0.017 (-0.33)	-1.257 (-5.35)	-7.434 (-3.57)	3.131 (11.43)	2.438 (1.20)
Housing-Price Residual		0.073 (1.01		2.534 (0.94)		-0.779 (-0.33)

Table 6: Incremental Modification of Specification						
Models	OLS	Fixed Effects Second Stage	IV Neighborhood Controls	IV with Housing Price Residual	IV for Housing Price Residual	Final IV Model
Labor Force Participant						
Employment-Access	0.006 (0.32)	0.027 (1.28)	0.620 (4.18)	0.601 (4.05)	1.461 (8.93)	0.308 (2.02)
Percent-Poverty	-0.528 (-20.62)	-0.523 (-30.66)	-1.674 (-13.65)	-1.665 (-13.52)	-2.250 (-16.82)	-1.055 (-8.29)
Percent-College Graduate	-0.099 (-15.01	-0.091 (-14.02)	-0.251 (-6.30)	-0.221 (-5.57)	-0.866 (-12.36)	-0.434 (-9.71)
Housing-Price Residual				-0.062 (-9.07)	0.696 (11.68)	0.514 (9.48)
Weeks Worked Last Year						
Employment-Access	-1.501 (-2.68)	1.011 (0.84)	43.157 (9.01	43.568 (9.10)	52.121 (9.72)	36.068 (7.42)
Percent-Poverty	-3.550 (-5.14)	-42.043 (-43.32)	-6.902 (-1.99)	-7.159 (-2.06)	-13.454 (-3.53)	6.027 (1.52)
Percent-College Graduate	1.264 (6.11)	1.251 (3.40)	-12.892 (-9.60)	-13.490 (-10.08)	-19.426 (-9.02)	-17.186 (-11.93)
Housing-Price Residual				1.301 (7.03)	7.243 (3.86)	11.685 (7.33)
Weekly Hours Last Year						
Employment-Access	1.769 (2.83)	-1.925 (-1.84)	23.570 (4.29)	23.066 (4.20)	31.185 (5.10)	19.156 (3.41)
Percent-Poverty	-11.176 (-15.03)	-32.24 (-37.98)	-22.723 (-5.36)	-22.413 (-5.28)	-28.302 (-5.90)	-14.709 (-3.18)
Percent-College Graduate	-2.397 (-12.26)	4.073 (12.64)	-11.251 (-7.83)	-10.521 (-7.31)	-16.80 (-6.69)	-13.929 (-8.62)
Housing-Price Residual				-1.601 (-8.70)	6.147 (2.87)	7.270 (3.76)

Table 7: Final IV Model for Alternative Sets of Neighborhood Controls

Models	No Neighborhood Controls	Poverty Only	Poverty and Employment Access	Plus Percent College Graduate	Plus Percent Minority Disadvantaged	Plus Percent not born in U.S.
Labor Force Participant						
Employment-Access			-0.166 (-1.11)	0.308 (2.02)	0.226 (1.47)	-0.039 (-0.24)
Percent-Poverty		-1.226 (-11.81)	-0.905 (-7.11)	-1.055 (-8.29)	-1.435 (-9.27)	-1.772 (-11.19)
Percent-College Graduate				-0.434 (-9.71)	-0.435 (-9.68)	-0.361 (-8.12)
Percent-Minority Disadvantage					0.186 (4.43)	0.053 (1.22)
Percent-Not-Born US						1.268 (7.49)
Housing-Price Residual	0.308 (9.90)	0.106 (3.45)	0.259 (6.24)	0.514 (9.48)	0.586 (10.59)	0.526 (9.27)
Weeks Worked Last Year						
Employment-Access			3.779 (0.71)	19.156 (3.41)	17.570 (3.11)	12.808 (2.20)
Percent-Poverty		-11.556 (-3.31)	-9.154 (-2.00)	-14.709 (-3.18)	-23.654 (-4.31)	-28.957 (-5.21)
Percent-College Graduate				-13.929 (-8.62)	-13.863 (-8.58)	-12.735 (-7.85)
Percent-Minority Disadvantage					4.20 (3.01	1.892 (1.28)
Percent-Not-Born US						21.906 (4.24)
Housing-Price Residual	-1.577 (-1.60)	-3.217 (-3.11)	0.303 (0.21)	7.270 (3.76)	8.628 (4.38)	7.794 (3.88)
Weekly Hours Last Year						
Employment-Access			20.378 (4.47)	36.068 (7.42)	34.843 (7.14)	28.495 (5.64)
Percent-Poverty		13.136 (4.39)	8.823 (2.25)	6.027 (1.52)	-0.474 (-0.10)	-7.005 (-1.43)
Percent-College Graduate				-17.186 (-11.93)	-17.202 (-11.88)	-15.898 (-10.97)
Percent-Minority Disadvantage					2.995 (2.35)	0.094 (0.07)
Percent-Not-Born US						28.035 (5.89)
Housing-Price Residual	-3.830 (-4.19)	-1.890 (-1.86)	0.454 (0.36)	11.685 (7.33)	12.837 (7.84)	12.136 (7.16)

Table 8: Relationship between Neighborhood and Capital Income						
Variables	Positive Capital Income		Capital Income Above $1,000		Capital Income Above $3,000	
	OLS	Final IV Model	OLS	Final IV Model	OLS	Final IV Model
Employment-Access	0.143 (3.82)	0.079 (0.41)	0.159 (4.70)	0.034 (0.19)	0.158 (5.19)	0.045 (0.26)
Percent-Poverty	-0.127 (-3.33)	-0.058 (-0.38)	-0.042 (-1.22)	0.065 (0.46)	0.012 (0.38)	0.152 (1.10)
Percent-College Graduate	0.084 (6.40)	0.051 (0.88)	0.043 (3.63)	.0002 (0.01	0.026 (2.40)	-0.034 (0.64)
Housing-Market Residual		-0.004 (-0.06)		0.038 (0.64)		0.058 (1.01

Table 9: Core and Supplemental Models of Labor Models Outcomes						
Core Labor Market Outcomes						
	Labor Force Participant		Weeks Worked Last Year		Average Hours per Week Last Year	
	OLS	Final IV Model	OLS	Final IV Model	OLS	Final IV Model
Employment-Access	0.006 (0.32)	0.308 (2.02)	1.769 (2.83)	19.156 (3.41)	-1.501 (-2.68)	36.068 (7.42)
Percent-Poverty	-0.528 (-20.62)	-1.055 (-8.29)	-11.176 (-15.03)	-14.709 (-3.18)	-3.550 (-5.14)	6.027 (1.52)
Percent-College Graduate	-0.099 (-15.01	-0.434 (-9.71)	-2.397 (-12.26)	-13.929 (-8.62)	1.264 (6.11)	-17.186 (-11.93)
Housing-Price Residual		0.514 (9.48)		7.270 (3.76)		11.685 (7.33)
Supplemental Labor Market Outcomes						
	Worked Last Year		Employment Last Week		Hours Worked Last Week	
	OLS	Final IV Model	OLS	Final IV Model	OLS	Final IV Model
Employment-Access	-0.004 (-0.25)	0.647 (3.98)	0.013 (1.17)	0.362 (3.10)	-1.324 (-2.20)	44.266 (8.34)
Percent-Poverty	-0.661 (-24.91)	-1.308 (-9.46)	-0.146 (-9.14)	-0.337 (-3.32)	-4.498 (-5.86)	1.906 (0.42)
Percent-College Graduate	-0.067 (-10.45)	-0.603 (-12.59)	0.009 (2.63)	-0.258 (-7.96)	0.646 (2.89)	-19.296 (-12.25)
Housing-Price Residual		0.704 (11.70)		0.201 (4.92)		12.920 (7.27)

Table 10: Incorporation of Neighborhood Housing Stock Controls						
Variables	Labor Force Participant		Weeks Worked Last Year		Average Hours per Week Last Year	
	Final IV Model	Control for Housing Stock	Final IV Model	Control for Housing Stock	Final IV Model	Control for Housing Stock
Employment-Access	0.308 (2.02)	0.591 (3.84)	19.156 (3.41)	20.941 (3.68)	36.068 (7.42)	38.235 (7.70)
Percent-Poverty	-1.055 (-8.29)	-1.079 (-8.60)	-14.709 (-3.18)	-14.580 (-3.18)	6.027 (1.52)	5.862 (1.49)
Percent-College Graduate	-0.434 (-9.71)	-0.633 (-12.36)	-13.929 (-8.62)	-15.650 (-8.97)	-17.186 (-11.93)	-19.122 (-11.92)
Housing-Stock Index		0.622 (10.21)		4.350 (2.25)		5.021 (2.58)
Housing-Market Residual	0.514 (9.48)	0.595 (11.43)	7.270 (3.76)	8.277 (4.48)	11.685 (7.33)	12.564 (8.14)