Does neighborhood income mix affect earnings of adults? New evidence from Sweden
Galster, G.; Andersson, R.; Musterd, S.; Kauppinen, T.M.

Published in:
Journal of Urban Economics

DOI:
10.1016/j.jue.2007.07.002

Citation for published version (APA):
Does neighborhood income mix affect earnings of adults? 
New evidence from Sweden

George Galster a,*, Roger Andersson b, Sako Muste dc, Timo M. Kauppinen d

a Department of Geography and Urban Planning, Wayne State University, Detroit, MI, USA
b Uppsala University, Uppsala, Sweden
c University of Amsterdam, Amsterdam, the Netherlands
d University of Helsinki, Helsinki, Finland

Received 1 June 2007; revised 5 July 2007
Available online 14 August 2007

Abstract
This paper contributes to the literature on obtaining unbiased estimates of neighborhood effects, explored in the context of a centralized social welfare state. We employ a longitudinal database comprised of all working age adults in metropolitan Sweden 1991–1999 to investigate the degree to which neighborhood income mix relates to subsequent labor incomes of adults and how this relationship varies by gender and employment status. We control for unobserved, time-invariant individual characteristics by estimating a first-difference equation of changes in average incomes between the 1991–1995 and 1996–1999 periods. We further control for unobserved time varying characteristics through an analysis of non-movers. These methods substantially reduce the magnitude of the apparent effect of neighborhood shares of low-, middle- and high-income males. Nevertheless, statistically and substantively significant neighborhood effects persist, though relationships are nonlinear and vary by gender and employment status. Males who are not fully employed appear most sensitive to neighborhood economic mix in all contexts.

© 2007 Elsevier Inc. All rights reserved.

1. Introduction

In the United States the geography of markets, institutions and governmental services is considerably varied in the degree to which opportunities for socioeconomic advancement are equally available regardless of residence (Galster and Killen, 1995). In many quarters there has been a growing concern about ecological inequalities at the local neighborhood scale, with particular attention paid to the constraints imposed by concentrated poverty environments (Wilson, 1987). Although during the 1990s the trend of increasing poverty concentration observed since 1970 in the US was reversed (Jargowsky, 2003), the long-term trend toward reducing economic diversity in American urban neighborhoods has persisted (Galster et al., 2005).

The pattern of how income groups are distributed across metropolitan space becomes of concern to urban economics and policy makers to the degree to which the economic mix of the local neighborhood independently and substantially affects the economic prospects of individual adults residing there. And on this point, there has been considerable scholarly debate, much of it methodological in nature, as we shall explain below. Central to this debate has been the degree to which analysts have...
been able to purge the effect of neighborhood economic mix from unobserved characteristics of individuals being studied that lead to various sorts of selection effects.

In this paper we contribute to this debate by providing two vehicles for controlling for characteristics of adults that likely affect their earnings and their neighborhood choices but are not directly observable. We employ these vehicles using data from the 1990s on an unusually large longitudinal sample: the adult metropolitan population of Sweden. We find statistically significant but small, non-linear effects from both low-income and high-income neighbors, which often vary by gender and employment status.

Our paper is organized as follows. Section II. begins with a brief overview of relevant empirical literature that emphasizes the challenge of selection/omitted variables bias, how prior work has attempted to overcome it, and how we grapple with it in this paper. We employ:

1. a first-difference specification as a control for unobserved, time-invariant individual characteristics and
2. a first-difference specification for only non-movers as a control for both time-invariant and time-varying individual characteristics that are unobserved.

We present our Swedish data and empirical approach in Section 3. Section 4. reports our earnings models for four strata of metropolitan adults distinguished by gender and employment status. We analyze consequences of varying the percentages of low-, middle- and high-income residents in neighborhoods, and probe for non-linear relationships in these effects. In Section 5 we consider which sorts of social interactions within neighborhoods are consistent with the observed relationships. In the final section we conclude and draw implications for housing and neighborhood policy.

2. Selection/omitted variables bias and measuring the magnitude of neighborhood effects

There has been a sizable literature devoted to measuring the independent magnitude of the effect of a neighborhood’s income composition on residents’ economic outcomes, employing multivariate statistical analyses on both cross-sectional and longitudinal databases of individuals. Most of this work has focused on the experience of neighborhood context as a child or adolescent providing a lagged labor market consequence; see: Payne (1987), Moffitt (1992), Corcoran et al. (1992), Haveman and Wolfe (1994, 1995), Gottschalk et al. (1994), Gottschalk (1996), Mayer (1997), Vartanian (1999a, 1999b), Pepper (2000), Ginther et al. (2000), Holloway and Mulherin (2004). Others, as in the current paper, have examined the relationship between neighborhood population characteristics and contemporaneous economic outcomes for adults; see O’Regan and Quigley (1996), Buck (2001), Musterd and Andersson (2005, 2006), Andersson et al. (2007), Bolster et al. (2004), Dawkins et al. (2005). In various national contexts these studies observed nontrivial partial correlations between the percentage of lower-income residents in a neighborhood and several measures of lagged or contemporaneous adult labor market performance. The findings have hardly been uniform, however; several have failed to find such a relationship: Ginther et al. (2000), McCulloch (2001), Musterd et al. (2003), and Drever (2004).

The accuracy of the neighborhood-labor market outcome relationships measured by these studies is subject to challenge, however, due to their failure to deal with selection/omitted variable bias (Ginther et al., 2000; Dietz, 2002). The basic issue is that adults have certain (unmeasured) motivations and skills related to their own (or their children’s) economic prospects and move to certain types of neighborhoods as a consequence of these attributes. Any observed relationship between neighborhood conditions and economic outcomes may therefore be biased because of this systematic spatial selection process, even if all the observable characteristics of parents are controlled (Manski, 1995). Flipped on its head, the problem can be formulated as omitted variables bias. Is the observed statistical relationship between outcomes and neighborhood composition indicative of neighborhood’s independent effect, or merely unmeasured (unobserved, uncontrolled) characteristics of adults that truly affected outcomes but also (spuriously, in the extreme) led to neighborhood choices as well?\footnote{The direction of the bias has been the subject of debate, with Jencks and Mayer (1990) and Tienda (1991) arguing that neighborhood impacts are biased upwards, and Brooks-Gunn et al. (1997) arguing the opposite.}

There have been several types of methodological responses to this challenge, each with distinct limitations. Again, no consensual set of findings has emerged:

- **Random Assignment Experiments**: Data are produced by an experimental design whereby households are randomly assigned to different neighborhoods, such as the Moving To Opportunity (MTO) demonstration. There appeared to be no important differences in adult labor market outcomes between
the experimental group who moved to low-poverty neighborhoods and control groups who remained in concentrated poverty neighborhoods (Goering et al., 2003; Orr et al., 2003).

- **Natural Quasi-Experiments:** Data are produced from idiosyncratic public policy initiatives (typically involving subsidized housing) that create exogenous variation in neighborhood environments for tenants. Studies employing this approach have reached diametrically opposed conclusions, even within the same national context (cf. Rosenbaum et al., 2002; Briggs, 1997, 1998; Oreopolis, 2003; Edin et al., 2003; Aslund and Fredricksson, 2005).

- **Fixed Effect Models based on Longitudinal Data:** Unobserved, time-invariant characteristics of individuals that may lead to both neighborhood selection and labor force outcomes are measured by individual dummy variables. Weinberg et al. (2004) find that hours worked and incomes are strongly related to neighborhood poverty rates in a non-linear fashion.

- **Difference models based on longitudinal data:** Unobserved, time-invariant characteristics are eliminated by measuring differences in between two periods. Bolster et al. (2004) find a British neighborhood deprivation index has statistically significant but small impact on individuals’ income growth.

- **Sibling Studies:** Longitudinal data on siblings is collected from surveys of households who reside in a variety of neighborhoods during the children’s upbringing but share unobserved parental characteristics in common (Aaronson, 1997, 1998; Plotnick and Hoffman, 1999). This method has not been employed to investigate impacts of neighborhoods on adult labor force outcomes.

- **Instrumental Variables for Neighborhood Characteristics:** Proxy variables for neighborhood characteristics are devised that only vary according to attributes exogenous to the household (Foster and McLanahan, 1996; Galster et al., in press). This method has not been employed to investigate contemporaneous impacts of neighborhoods on adult labor force outcomes.

We employ the difference model approach here because it directly eliminates time-invariant characteristics of adults that typically are unobservable and uncontrolled. We also employ this approach for a subset of individuals who do not move during the analysis period. Because changes observed in their neighborhood will be uncorrelated with their unobserved individual characteristics (both time-varying and time-invariant) the estimated relationship should be unbiased. To our knowledge this approach has not been tried previously.

### 3. Data and empirical model

#### 3.1. The Swedish data files

The variables we employ are constructed from data contained in the Statistics Sweden *Louise* files, which are produced annually. These files contain a large amount of information on all individuals age 15 and above and represent compilations of data assembled from a range of statistical registers (income, education, labor market, and population). We have merged selected information about individuals from annual *Louise* files to create a longitudinal database 1991–1999 for all individuals present in Sweden in 1991. Here we analyze only those residing in one of Sweden’s three large metropolitan areas: Stockholm, Gothenburg, and Malmö, so we can ensure a meaningfully consistent concept of neighborhood (explained below). We confine our sample to working-age individuals (ages 20–60 in 1996) who were residents of Sweden in 1991 and present each year 1991–1999, producing an analysis sample size of 1,689,704. We categorize individuals according to whether they work full time (defined by us as more than 152 days per year for males or 144 days per year for females) or not.

We emphasize that our dataset includes observations of virtually the entire metropolitan population within the desired adult age and residency range, not a sample. Thus, the *t*-statistics we present below should not be interpreted as guides for prospective errors involving inferences from a sample to the larger population. Rather, they provide a means of assessing the reliability of estimated coefficients as parameters for the underlying gender-specific income function for individuals in Swedish metropolitan areas, given potential functional misspecifications and measurement errors in variables.

#### 3.2. Our model for explaining individual incomes

Our outcome of interest is the average annual income from work (measured in Swedish kronor, SEK;
$1 = 7.8$ SEK) during a four year period. Since this indicator encapsulates educational credentials, labor force participation, employment regularity, and hourly compensation, we believe it to be the most comprehensive single measure of an individual’s economic worth. Descriptive statistics for this outcome variable are presented in Table 1, stratified by gender, for both periods used in our analysis. Average incomes for both genders grew substantially between the 1991–1994 and 1996–1999 periods in metropolitan Sweden: 24.7% for males and 22.3% for females. In the earlier period males earned around 40.5% more than females, on average; in the later period this had risen to 43.1%.

We model the average income from work during period $t + 1$ to $t + 4$ ($I_{t+4}$) for individual $i$ residing in neighborhood $j$ in local labor market area $k$ as:

\[
\ln(I_{ijkt-4}) = \alpha + \beta[P_{it+1}] + \gamma[P_i] + \phi[UP_{it+1}]
+ \partial[UP_i] + \theta[N_{ij}]
+ \mu[L_{kt4}] + \varepsilon
\] (1)

where:

$\{P_{t+1}\} =$ observed personal characteristics that can vary over time (e.g., marital or fertility status, educational attainment);

$\{P\} =$ observed personal characteristics that do not vary over time (e.g., year and country of birth);

$\{UP_{t+1}\} =$ unobserved personal characteristics that can vary over time (e.g., psychological states, interpersonal networks);

$\{UP\} =$ unobserved personal characteristics that do not vary over time (e.g., IQ, prior experiences);

$\{N_{ij}\} =$ observed characteristics of neighborhood where individual resides at time $t$;

$\{L_{kt}\} =$ observed characteristics of local labor market area in which individual resides during $t + 1$ to $t + 4$ (e.g., mean earnings);

$\varepsilon =$ a random error term with statistical properties discussed below.

Equation (1) is estimated separately for males and females. We model ($I_{t+4}$) as a function of residence during $t$, realizing that people can and do move during the following period. This lagged specification is intentional so we can keep causation clear, since we know that changes in income can lead to changes in residence. Our point is to test whether initial residential context has any relationship to subsequent income flows, regardless of whether those flows lead to residential mobility or not. Details of variable specifications follow.

In this study we operationalize “neighborhood” as the area delineated by a “SAMS” defined by Statistics Sweden. The SAMS classification scheme is designed to identify relatively homogeneous areas by taking into account housing type, tenure and construction period; as such they are comparable to a US census tract. We confine our analyses to the three Swedish metropolitan areas so that geographic scale of neighborhood can be made more comparable across individuals being analyzed. We emphasize that using such a large scale of neighborhood likely has the effect of reducing the measured neighborhood effects, inasmuch as other studies using European data have consistently found stronger effects at smaller spatial scales (Buck, 2001; Bolster et al., 2004).

We focus on the income mix of neighborhood as the $\{N\}$ variable of importance for three reasons. First, this is the aspect of neighborhood that has been the focus of the scholarly literature beginning with the “concentrated poverty” thesis of Wilson (1987). Second, this dimension has been the focal point of several public policy initiatives in both the US and Western Europe. Third, an earlier study using similar Swedish data found that initial neighborhood income mix was more strongly correlated with subsequent levels of individual incomes than neighborhood mix defined by education, ethnicity, or housing tenure (Andersson et al., 2007). As our measure of neighborhood income mix we specify the proportion of working age males in the lowest 30% of the (sample) male income distribution and that proportion in the highest 30% of the distribution; the middle 40% becomes the excluded reference category. For brevity we will subsequently refer to these groups as “lower-income,” “middle-income,” and “higher-income” neighbors. We adopt this convention because of the longstanding precedent in the literature investigating the external consequences of

---

3 Formally, income from work is computed here as the sum of: cash salary payments, income from active businesses, and tax-based benefits that employees accrue as terms of their employment (sick or parental leave, work-related injury or illness compensation, daily payments for temporary military service, or giving assistance to a handicapped relative).

4 The log-linear transformation not only is appropriate given the positive skew of the income distribution, but also has sound grounding in economic theory, implicitly suggesting that income is a multiplicative (not additive) function of personal, neighborhood, and labor market characteristics.

5 There remains some unavoidable inter-urban variation in SAMS scale nevertheless. At the extremes, the average SAMS in Gothenburg has a population of about 500 but in Stockholm it contains over ten times as many people.
Table 1
Descriptive statistics of all variables, by gender and period

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Males</td>
<td>Females</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Annual mean labor income (100 SWE kroner)</td>
<td>1743.59</td>
<td>1335.71</td>
</tr>
<tr>
<td>Neighborhood variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion in lowest 3 male income deciles</td>
<td>0.30</td>
<td>0.12</td>
</tr>
<tr>
<td>Proportion in highest 3 male income deciles</td>
<td>0.36</td>
<td>0.15</td>
</tr>
<tr>
<td>Control Variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td># years with pre-retirement benefits</td>
<td>0.15</td>
<td>0.71</td>
</tr>
<tr>
<td># cumulative child-years, for children under 7</td>
<td>1.24</td>
<td>2.42</td>
</tr>
<tr>
<td># years enrolled in school during period</td>
<td>0.28</td>
<td>0.84</td>
</tr>
<tr>
<td># years with parental leave benefits</td>
<td>0.69</td>
<td>1.24</td>
</tr>
<tr>
<td># years with sick leave benefits</td>
<td>1.08</td>
<td>1.10</td>
</tr>
<tr>
<td>Immigrants w/ &lt; 5 years in Sweden (1 = yes)</td>
<td>0.05</td>
<td>0.21</td>
</tr>
<tr>
<td>1–11 years education (1 = yes)</td>
<td>0.09</td>
<td>0.29</td>
</tr>
<tr>
<td>11–12 years education (1 = yes)</td>
<td>0.42</td>
<td>0.49</td>
</tr>
<tr>
<td>13–14 years of education (1 = yes)</td>
<td>0.13</td>
<td>0.33</td>
</tr>
<tr>
<td>15+ years of education (1 = yes)</td>
<td>0.15</td>
<td>0.36</td>
</tr>
<tr>
<td>Civil status: couple (1 = yes)</td>
<td>0.57</td>
<td>0.50</td>
</tr>
<tr>
<td>Mean income in local labor market (100 SWE kroner)</td>
<td>1412.67</td>
<td>89.49</td>
</tr>
<tr>
<td>N</td>
<td>847162</td>
<td>842542</td>
</tr>
</tbody>
</table>

* Omitted category = no formal education.

both the most- and least-advantaged segments of the neighborhood’s population (Jencks and Mayer, 1990; Brooks-Gunn et al., 1997). In our database we only observe these neighborhood conditions at two points.

We operationalize the observed personal characteristics of individuals $[P_t]$ and $[P]$ with a set of variables describing their demographic and household characteristics, educational attainments, nativity and immigrant status, and features of their employment during the period that will affect their income but are likely not related to neighborhood context (such as parental leave or attending school). We operationalize $[L_t]$ with the mean labor earnings for the local labor market (an area somewhat smaller than a metropolitan area) in which the individual resided during the period in question. Descriptive statistics for all variables for the two periods to be analyzed are presented for males and females in Table 1.

3.3. Strategy for estimating the potential causal effect of neighborhood income mix

As noted above, the principal challenge in quantitative neighborhood effects research is avoiding biased estimates of $\theta$ arising from failure to control for $[UP]$ and $[UP]$ in (1). We address this challenge in two ways. The first is to exploit the fact that we have neighborhood conditions, personal characteristics, and multi-year average incomes measured at two points in time. By differencing (1) between these two points (let this time difference be denoted $f$) we obtain:

$$\Delta^f \ln(I_{ijk}) = \beta \Delta^f P_{it} + \phi \Delta^f UP_{it} + \theta \Delta^f N_{jt} + \mu \Delta^f L_{kt} + \Delta^f \epsilon.$$ (2)

Note that differencing removes the troubling unobserved, time-invariant individual characteristics $[UP]$, but not the time-varying ones $[UP_{it}]$.

The second strategy is to exploit our ability to distinguish in estimating (2) individuals who do not change neighborhoods between the first and second observations. For these non-movers the changes we observe in their neighborhoods’ income mix must be exogenous, and therefore uncorrelated with any remaining unobserved individual time-varying characteristics. Thus, even though we cannot control $\Delta^f UP_{it}$ in (2), $\theta$ should be unbiased. There are three caveats associated with this strategy. First, those who do not move may not be representative of the larger population. Second, they are likely to experience smaller variation in neighborhood income mix; indeed, in our data the means and standard deviations of the changes in neighborhood shares of low-income and high-income males observed for non-movers are less than one third the magnitudes ob-
served for movers. Moreover, if the effect of neighborhood income mix on an individual’s subsequent income-earning capacity is cumulative, those who have resided in their initial location at least for four years should evince a larger $\theta$ than those who have moved in more recently.

4. The consequences of neighborhood income mix in Swedish metropolitan areas: empirical results

4.1. Baseline results

Parameter values for the two neighborhood income mix variables obtained by estimating (1) via OLS and employment status are presented in the upper panel of Table 2, stratified by gender and employment status. These values for $\theta$ may be thought of as those that are manifested without any attempts to overcome selection/omitted variables bias. They appear nontrivial and are highly statistically significant, suggesting that the distribution of low-, middle- and high-income males in the neighborhood correlates with residents’ subsequent income. Note that the signs of the coefficients of neighborhood group $X$ in this and succeeding tables should be interpreted as the differences in otherwise-comparable individuals’ incomes associated with residence in neighborhoods that differ only in their share of $X$ and middle-income residents, holding the share of the remaining group constant. All results in Table 2 are produced controlling for the full set of 1995 characteristics listed in Table 1.

The key question remains, however. To what extent are these baseline relationships the product of independent causation by neighborhood income mix or mere spurious correlation due to omitted variables that affect neighborhood selection and income? The next two sections address the question.

4.2. The effect of differencing

The second panel of Table 2 shows the consequences for the parameters of the neighborhood income mix variables when they are estimated from our difference specification (2), for full results for all control variables, see Appendix Tables A and B. In all cases the estimated standard errors remain essentially unchanged but the coefficients shrink dramatically, suggesting that indeed unobserved, time-invariant personal characteristics are important to control here to reduce bias in $\theta$. The differencing model yields greater reductions in the coefficients of the percentage of low-income neighbors than in the coefficients of the percentage of high-income neighbors, as would be expected if there were stronger selection away from places with predominantly lower-income groups than toward places with predominantly higher-income places (cf. Ginther et al., 2000). The results of the difference model show no general patterns about comparative neighborhood effect sizes across the various gender or employment status categories.

In the difference model, the share of high-income neighbors is positively associated with higher income growth subsequently in a statistically significant way (with the exception of males employed less than full-time). By contrast, for all but one stratum the coefficient of the percentage low-income neighbors fails to achieve statistical significance; the one exception is the positive coefficient for fully employed males. This result initially appears puzzling because it implies that lower-income neighbors produce superior outcomes for this stratum compared to middle-income ones. As we shall see below, this relationship disappears for those who do not move, so we believe this finding is the result of particular selections related to changes in $[\theta_{P1}]$ during our analysis frame. One possibility is that there were substantial numbers of “gentrifiers” who had unobserved changes in their attributes that both increased their earnings and systematically led them to move into up-and-coming but relatively deprived neighborhoods in Swedish central cities. Another possibility is that those with unobserved changes in attributes leading to lower incomes were able to access higher-income neighborhoods by entering the public housing system. A final possibility is that Swedish national policy initiatives designed to increase the tenure mix of large social housing estates may have financially induced some younger males with rising incomes to buy first homes or cooperative apartments in neighborhoods with higher percentages of low-income neighbors than they normally would have chosen.

4.3. The effect of restriction to non-movers

We reestimated the difference model (2) separately for non-movers in each of the gender/employment categories; the results are presented in the bottom panel of Table 2. Predicted changes in coefficient magnitudes that result from this procedure are unclear, because of

6 The mean changes in the proportion of low-(high-) income neighbors observed in the full sample, mover sample, and non-mover samples are: 0.064 (0.070); 0.109 (0.122); 0.034 (0.035), respectively. The corresponding standard deviations are: 0.081 (0.081); 0.107 (0.104); 0.034 (0.028).
the potentially offsetting impacts of controlling for selection on \([UP_r]\) and the exposure to neighborhood for a longer duration. Here the impact appears to be dominated by the latter, as the magnitudes are generally larger for the non-mover stratum. However, in the case of the coefficients for the percentage of high-income neighbors, only for fully employed males does the (expected, positive) coefficient retain statistical significance. In the case of the coefficients for the percentage of low-income neighbors, only for fully employed males and fully employed females does the (expected, negative) coefficient retain statistical significance. The aforementioned puzzling changes for fully employed males produced for the low-income share parameter by the difference model disappear here. The apparent negative impact of low-income neighbors on males who are not fully employed has the largest effect size here; the positive impact of high-income neighbors on fully employed males is half as large and the negative impact of low-income neighbors on fully employed females is one fifth as large. Before exploring further the economic significance of these parameters, however, we first explore potential non-linearities in the neighborhood income mix–individual income relationship, beyond those imposed by the semi-log function.

### 4.4. Explorations of non-linear relationships

There are a variety of theoretical reasons why the underlying social interactions potentially yielding neighborhood effects operate in non-linear fashion (Galster, 2002). Several prior studies have found non-linear effects of the share of disadvantaged neighbors (Varianian, 1999a, 1999b; Weinberg et al., 2004) and/or advantaged neighbors (Crane, 1991; Chase-Lansdale et al., 1999a, 1999b) on individuals’ labor market performance. Given the importance of such non-linearities in forming conclusions about the appropriateness of engaging in policies aimed at deconcentrating disadvantaged populations (Galster, 2002, 2007), we probe for them here. Specifically, we specify interaction terms that allow both of the neighborhood income mix variables to take on different coefficients depending on whether the neighborhood started the period above or below the median of that variable. This specification allows us to test whether there is a difference in the marginal impact of changing neighborhood income mixes depending upon the range of their starting values.

The results of this exercise undertaken for the non-mover stratum are reported in Table 3. Comparisons of the estimated parameters in the corresponding cells in the “above median” and “below median” panels of
Table 3 provide strong evidence of substantial, though complex and varied, nonlinearities. Ten of the sixteen coefficients across the gender/employment strata prove statistically significant.

The first notable nonlinear finding is that, by far, the strongest (in terms of size of coefficient) neighborhood income mix effect occurs when males who are not fully employed reside in neighborhoods where both shares of low- and high-income neighbors are below their respective medians. In these circumstances, replacing middle-income neighbors with either low-income and/or high-income neighbors reduces their subsequent incomes. The same is true (but for considerably smaller magnitudes) for females who are not fully employed, at least in terms of replacing middle-income with low-income neighbors. This is suggestive that Swedes who are not fully employed benefit from residing in a place with predominantly middle-income neighbors; as this critical mass is eroded by replacements from either other group, their income-earning prospects deteriorate. It may be that middle-income neighbors provide those who are not fully employed with “bridging social capital,” networks that contain valuable economic information (unlike the case with contacts with low-income neighbors) and are not impeded by excessive social distance (as would be the case with high-income neighbors).

Interestingly, in the case of not fully employed males (but not females), marginal increases in low-income neighbors replacing middle-income ones continue to have negative (though smaller) impacts in places where the share of the former initially exceeded its median, which is consistent with negative role modeling, norm and peer effects emanating from the low-income group. But the apparent impact on males who are not employed from replacing middle-income neighbors with high-income ones is completely different (i.e., positive) in neighborhoods that started with above-median high-income shares. This implies that removing the beneficial bridging social capital that the middle-income neighbors provide is a net loss to males who are not fully employed up until the point where the high-income group becomes predominant in the area, whereupon some different, net-positive externality starts to occur on the margin. We cannot be sure what this is, but it may be that the growth of the neighborhood retail and commercial sectors associated with the predominance of the high-income group may enhance local income-generating opportunities for male neighbors.

A second intriguing nonlinear finding is that for both males and females who are employed full time there are diminishing returns to the share of high-income neighbors. In neighborhoods initially below the median of this share, subsequent increases yield gains in incomes for these individuals, consistent with a causal mechanism of inter-neighbor networking whereby the first few high-income people living nearby help fully employed residents find better-paying jobs than they otherwise could have. However, as the high-income share starts to predominate in the neighborhood, this positive marginal impact is attenuated (and even turns slightly negative in the case of females). We do not have a plausible explanation for this result.

Table 3
Regression results for neighborhood income mix variables, nonlinear models

<table>
<thead>
<tr>
<th>Neighborhood income mix variables</th>
<th>% Low income males (deciles 1–3)</th>
<th>b</th>
<th>s.e.</th>
<th>beta</th>
<th>p</th>
<th>% High income males (deciles 8–10)</th>
<th>b</th>
<th>s.e.</th>
<th>beta</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above median in 1991</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–152 employment days in 1995</td>
<td>0.742</td>
<td>0.210</td>
<td>-0.013</td>
<td>&lt; 0.001</td>
<td>1.113</td>
<td>0.390</td>
<td>0.010</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;152 employment days in 1995</td>
<td>0.035</td>
<td>0.042</td>
<td>0.001</td>
<td>0.409</td>
<td>0.247</td>
<td>0.049</td>
<td>0.008</td>
<td>&lt; 0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–144 employment days in 1995</td>
<td>0.207</td>
<td>0.191</td>
<td>0.004</td>
<td>0.278</td>
<td>0.117</td>
<td>0.327</td>
<td>0.001</td>
<td>0.720</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;144 employment days in 1995</td>
<td>-0.094</td>
<td>0.043</td>
<td>-0.003</td>
<td>0.030</td>
<td>-0.144</td>
<td>0.050</td>
<td>-0.004</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below median in 1991</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–152 employment days in 1995</td>
<td>-1.509</td>
<td>0.438</td>
<td>-0.012</td>
<td>&lt; 0.001</td>
<td>-1.429</td>
<td>0.3391</td>
<td>-0.0154</td>
<td>&lt; 0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;152 employment days in 1995</td>
<td>-0.040</td>
<td>0.058</td>
<td>0.001</td>
<td>0.491</td>
<td>0.403</td>
<td>0.0555</td>
<td>-0.0115</td>
<td>&lt; 0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–144 employment days in 1995</td>
<td>-0.814</td>
<td>0.373</td>
<td>-0.007</td>
<td>0.029</td>
<td>0.412</td>
<td>0.3014</td>
<td>0.0046</td>
<td>0.172</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;144 employment days in 1995</td>
<td>-0.062</td>
<td>0.058</td>
<td>-0.002</td>
<td>0.285</td>
<td>0.335</td>
<td>0.0574</td>
<td>0.0092</td>
<td>&lt; 0.01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Estimates are based on non-mover sample.
4.5. Assessing the economic significance of neighborhood income mix

Our various tests consistently indicate that there likely is an independent, causal impact of neighborhood income mix on some individuals’ subsequent incomes in Sweden, at least in some neighborhood contexts. We assess the economic significance of these findings as follows. Based on the statistically significant parameters presented in Table 3, Table 4 portrays the predicted percentage change in incomes of the “median” male and female (non-movers) in fully and not fully employed categories, if they were to reside in various hypothetical neighborhood environments.7 For simplicity of presentation, we consider changes in mix starting from two alternative baselines: both the proportions of low-income and high-income neighbors are above or below their respective median values at the beginning of the period.8 Heuristically, this corresponds to two alternatives where

---

7 “Median person” means a person whose predicted income in a neighborhood with median values for the percentages of both low- and high-income males (in the whole data set) is the median income in his/her gender and employment category. At the beginning of our analysis period, the median proportion of low-(high-) income males = 0.283 (0.350).

8 Simulations based on the remaining two permutations are available from the first author upon request.

---
be 20 percent greater. Other hypothetical comparisons can easily be made with the estimates presented in Table 4.

5. Implications for mechanisms of social interaction in Swedish metropolitan neighborhoods

We gain some fascinating insights by placing the foregoing results in the context of extant theories about social interactions within neighborhoods. Our evidence suggests that different sorts of neighborhood effect mechanisms (negative role modeling/peer effects, network effects) may be operating upon and among different groups.

In the case of men who work full time, living in neighborhoods with ever-higher shares of high-income neighbors enhances their subsequent incomes. We think it is likely that having higher-income neighbors is considerably superior to having either other group as neighbors because they provide more valuable information about higher-paying job opportunities to their already-employed (presumably high-income) neighbors. This network mechanism is more plausible than an alternative, positive role model mechanism, because presumably the fully employed group in question does not need such role models to enhance their earnings.

For males who were not employed full time at the beginning of our analysis period (and who typically have lower-incomes themselves), it is the neighborhood with the highest possible share of middle-income neighbors that is most conducive to their earning more. The fact that even a few low-income neighbors erode these benefits suggests that a negative role modeling/peer effect may be transpiring here. It does not appear that the epidemic/social norm model of interaction is consistent with these data, for no minimum threshold of low-income neighbors was observed past which their negative impacts began. Replacing middle-income with high-income neighbors also has negative impacts on these males, implying that the former provide important employment-related information networks but the latter do not, perhaps because many of the jobs that they might provide information about are beyond the qualifications of low-income neighbors or because the social distance between the groups is too great for networks to develop. Apparently, unless middle-income neighbors constitute a dominant share there will be an insufficiently diverse, approachable, resource-laden network for those with weaker labor force attachments to access. We find little support for the hypothesis of positive role models for males here, because such would imply no distinctions between shares of middle- and high-income neighbors under the assumption that both provide comparable role models.9

This implication of valuable intra-neighborhood job information networks between low- and middle-income (but not low- and high-income) neighbor groups in Sweden generally comports well with the US sociological literature suggesting that if the social class gap is too wide there will be little meaningful social interaction among neighbors (Kleit 2001a, 2001b, 2002, 2005; Clampet-Lundquist, 2004). However, there is some literature suggesting that even relatively narrow neighborhood class differences also may be difficult to bridge with resource-rich networks for the poor (Briggs 1997, 1998; Schill, 1997), though apparently that is not the case with the lower-income Swedish individuals here.

6. Conclusions and implications

In this paper we have attempted to contribute to the literature on obtaining unbiased estimates of the size of neighborhood effects. We employ an unusually large longitudinal database comprised of all working age adults in metropolitan Swedish neighborhoods during the period 1991–1999 to assess the determinants of earnings with a great degree of precision. This database allows us to estimate difference models for controlling time-invariant and, when applied to non-movers, time-varying unobserved individual characteristics that may affect both neighborhood selection and earnings. When these models are employed the magnitude of the apparent neighborhood effect is strongly attenuated. Nevertheless, the models reveal statistically and substantively significant effects, often nonlinear in nature, of neighborhood shares of low-, middle-, and high-income neighbors on subsequent earnings.

The robust results indicate that:

1. For males who are not employed full time, middle-income neighbors have a positive marginal impact (relative to either high- or low-income neighbors);
2. For males who are employed full-time, high-income neighbors (relative to any mix of the other two groups) have a positive marginal impact; and

9 Of course, if low-income residents view higher-income neighbors as an inappropriate reference group because they are beyond the feasible realm of emulation, middle-income neighbors might be used as the source of role modeling. We are indebted to an anonymous referee for pointing this out.
(3) the same is true for fully employed females when the share of high-income neighbors is relatively small in the neighborhood.

We believe that this evidence is consistent with the view that, for males who are not fully employed, low-income neighbors provide negative role models and middle-income (but not high-income) neighbors provide access to networks with valuable employment-related information. For those already fully employed, high-income neighbors probably are valuable because they provide access to networks with information about opportunities for more lucrative employment. These results comport with a growing literature on job networks (Ioannides and Loury, 2004).

The observed importance of neighborhood income mix in shaping earnings of its residents (particularly males not employed full time) may come as a surprise to those who know of Sweden’s longstanding attempts to reduce cross-sectional variation in this indicator. Indeed, some have suggested that neighborhood effects might be small in any event, given the myriad of income support structures and high-quality public and private institutions available to all in a centralized social welfare state like Sweden, regardless of location (Friedrichs, 2002; Musterd, 2002). That we are able to identify such neighborhood effects attests to the statistical power provided by our large number of observations. That they exist to a nontrivial degree suggests that even a comprehensive welfare state cannot overcome the function of role modeling and interpersonal networks in shaping economic opportunities.

In political terms, our findings provide support to those who argue that housing policies need to alter the neighborhood economic mix to help those with marginal labor force attachments. Providing opportunities for males who work less than full time to reside with middle-income neighbors would likely result in substantial income gains for them, if the Swedish results can be generalized. Whether or not governmental authorities can find effective measures that might bring this about—without interfering with democratic ideals of peoples’ free residential choices—is a distinct and important question beyond the scope of this paper.

Of course, our Swedish findings should not necessarily be generalized to the US context. In this paper we have defined “low-income neighbors” much more expansively than the common US measure “percent living below the poverty line.” Moreover, the comparatively limited spatial variation in neighborhood income diversity in Sweden, coupled with its social welfare system, likely implies that the nature and function of neighborhood social interactions in shaping opportunities is different between the two nations. Yet, we think it reasonable to speculate that, were comparable data available in the US, replicating our analysis might well reveal larger neighborhood effects than we report here.

Acknowledgments

The authors wish to thank Jan Brueckner and anonymous referees for their invaluable suggestions on an earlier draft. Nina Butler provided production assistance.

Appendix Table A
Regression results for difference model: males

<table>
<thead>
<tr>
<th></th>
<th>0–152 employment days in 1995</th>
<th>&gt;152 employment days in 1995</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>s.e.</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.5826</td>
<td>0.0107</td>
</tr>
<tr>
<td>Change in # years with pre-retirement benefits</td>
<td>-0.8623</td>
<td>0.0074</td>
</tr>
<tr>
<td>Change in # child-years (children under 7)</td>
<td>-0.3052</td>
<td>0.0038</td>
</tr>
<tr>
<td>Change in # years studying</td>
<td>-0.3052</td>
<td>0.0038</td>
</tr>
<tr>
<td>Change in # years with parental leave</td>
<td>0.3740</td>
<td>0.0084</td>
</tr>
<tr>
<td>Change in # years with sick leave</td>
<td>0.3540</td>
<td>0.0053</td>
</tr>
<tr>
<td>Recent immigrant in 1991, not in 1995</td>
<td>0.0291</td>
<td>0.0226</td>
</tr>
<tr>
<td>Education rose from low level*</td>
<td>1.0678</td>
<td>0.0399</td>
</tr>
<tr>
<td>Education rose from medium level*</td>
<td>0.9853</td>
<td>0.0394</td>
</tr>
<tr>
<td>Civil status changed from couple to single**</td>
<td>0.2283</td>
<td>0.0230</td>
</tr>
<tr>
<td>Civil status changed from single to couple**</td>
<td>0.4604</td>
<td>0.0274</td>
</tr>
<tr>
<td>Change in mean income in local labor market</td>
<td>0.0022</td>
<td>0.0001</td>
</tr>
<tr>
<td>Change in % in lowest 3 male income deciles</td>
<td>-0.1089</td>
<td>0.0770</td>
</tr>
<tr>
<td>Change in % in highest 3 male income deciles</td>
<td>-0.0998</td>
<td>0.0938</td>
</tr>
<tr>
<td>R-square</td>
<td>0.167</td>
<td>0.191</td>
</tr>
</tbody>
</table>

* Excluded category = no change in education credentials during period.
** Excluded category = no change in civil status during period.
Appendix Table B
Regression results for difference model: females

<table>
<thead>
<tr>
<th></th>
<th>0–144 employment days in 1995</th>
<th>&gt;144 employment days in 1995</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>s.e.</td>
</tr>
<tr>
<td>Constant</td>
<td>−0.4072</td>
<td>0.0096</td>
</tr>
<tr>
<td>Change in # years with pre-retirement benefits</td>
<td>−0.7634</td>
<td>0.0071</td>
</tr>
<tr>
<td>Change in # child-years (children under 7)</td>
<td>−0.1588</td>
<td>0.0027</td>
</tr>
<tr>
<td>Change in # years studying</td>
<td>−0.1647</td>
<td>0.0055</td>
</tr>
<tr>
<td>Change in # years with parental leave</td>
<td>0.3859</td>
<td>0.0052</td>
</tr>
<tr>
<td>Change in # years with sick leave</td>
<td>0.4250</td>
<td>0.0047</td>
</tr>
<tr>
<td>Recent immigrant in 1991, not in 1995</td>
<td>0.4924</td>
<td>0.0209</td>
</tr>
<tr>
<td>Education rose from low level*</td>
<td>0.9502</td>
<td>0.0328</td>
</tr>
<tr>
<td>Education rose from medium level*</td>
<td>0.7436</td>
<td>0.0340</td>
</tr>
<tr>
<td>Civil status changed from couple to single**</td>
<td>0.1806</td>
<td>0.0212</td>
</tr>
<tr>
<td>Civil status changed from single to couple **</td>
<td>0.3916</td>
<td>0.0239</td>
</tr>
<tr>
<td>Change in mean income in local labor market</td>
<td>0.0018</td>
<td>0.0001</td>
</tr>
<tr>
<td>Change in % in lowest 3 male income deciles</td>
<td>0.1191</td>
<td>0.0785</td>
</tr>
<tr>
<td>Change in % in highest 3 male income deciles</td>
<td>0.2146</td>
<td>0.0858</td>
</tr>
<tr>
<td>R square</td>
<td>0.202</td>
<td></td>
</tr>
</tbody>
</table>

* Excluded category = no change in education credentials during period.
** Excluded category = no change in civil status during period.

References


