Housing Studies,
Vol. 22, No. 5, 637–660, September 2007

What Mix Matters? Exploring the Relationships between Individuals' Incomes and Different Measures of their Neighbourhood Context

ROGER ANDERSSON*, SAKO MUSTERD**, GEORGE GALSTER† & TIMO M. KAUPPINEN‡

*Uppsala University, Sweden, **University of Amsterdam, The Netherlands, †Wayne State University, Detroit, MI, USA, ‡University of Helsinki, Finland

(Received October 2005; revised 1 May 2007)

ABSTRACT *There is substantial interest among policy makers in both Western Europe and North America in reducing concentrations of disadvantaged households through initiatives to enhance the 'social mix' of neighbourhoods. However, there is little consideration or understanding with regard to which mix of household characteristics matters most in influencing the socio-economic outcomes for individual residents. This paper explores the degree to which a wide variety of 1995 neighbourhood conditions in Sweden are statistically related to earnings for all adult metropolitan and non-metropolitan men and women during the 1996–99 period, controlling for a wide variety of personal characteristics. The paper finds that the extremes of the neighbourhood income distribution, operationalized by the percentages of adult males with earnings in the lowest 30th and the highest 30th percentiles, hold greater explanatory power than domains of household mix related to education, ethnicity or housing tenure. Separating the effects of having substantial shares of low and high income neighbours, it is found that it is the presence of the former that means most for the incomes of metropolitan and non-metropolitan men and women, with the largest effects for metropolitan men.*

KEY WORDS: Neighbourhood effects, household mix, longitudinal, Sweden

Research and Policy Context

Scholarly interest in the degree to which neighbourhood conditions affect the life chances of children, youth and adults has skyrocketed over the last 15 years, as evidenced by the upsurge of research papers, special issues of scholarly journals and reviews of the literature.[1] Although there seems to be an emerging consensus in the US that

Correspondence Address: Roger Andersson, Department of Social and Economic Geography, Institute for Housing and Urban Research, Box 256, Uppsala University, Uppsala 751 05, Sweden. Email: Roger.Andersson@ibf.uu.se

ISSN 0267-3037 Print/1466-1810 Online/07/050637–24 © 2007 Taylor & Francis
DOI: 10.1080/02673030701474602

neighbourhood indicators typically do have non-trivial relationships with a range of outcomes related to childhood cognitive development, academic achievement and credential attainment, teen fertility and labour market consequences as adults, there remains debate about which precise indicators of neighbourhood conditions are the most predictive. There is even less consensus in Western European scholarship that *any* types of neighbourhood effects are prominent, let alone which indicator is most powerful (cf. Andersen, 2003; Kearns, 2002; Musterd, 2002; Musterd *et al.*, 2003; Musterd & Andersson, 2005, 2006; Ostendorf *et al.*, 2001; Van der Klaauw & Van Ours, 2003).

However, this lack of scholarly consensus has not slowed neighbourhood diversification policy initiatives on either side of the Atlantic (Friedrichs *et al.*, 2003; Galster & Zobel, 1998; Musterd *et al.*, 2003; Murie & Musterd, 2004).[2] A common guiding principle appears to be that concentrations of lower-income, disproportionately ethnic minority (typically immigrant in Europe) households need to be replaced by mixed residential environments that will provide superior social opportunities for the disadvantaged group. Although important differences remain between US and Western European approaches (cf. Goetz, 2002; Johnson *et al.*, 2002; Kearns, 2002; Musterd, 2002), this goal of enhancing the mix of social groups at the neighbourhood level has been sought through two common programmatic strategies. First, in Europe there have been large-scale investments aimed at restructuring large, homogeneous, post-war neighbourhoods and housing estates (through selective demolition, infill construction and the sale of social or public housing) so that they contain a greater diversity of housing types by price range and tenure (Murie *et al.*, 2003). Although smaller in scale, the redevelopment of US public housing developments as mixed-income complexes through the HOPE VI programme is strategically comparable. Second, it is now required in several jurisdictions in both Europe and the US that new, larger-scale residential developments must set aside a minimum share of the dwelling units for social (subsidized) housing.

What typically has been left implicit in this set of policy initiatives is what aspect of neighbourhood mix is deemed crucial for expanding opportunities. Is the domain of mix: socio-economic? National origin or ethnicity? Housing type and tenure? Within a domain, is it the percentage of a certain 'disadvantaged' group or the percentage of 'advantaged' groups that is more important? Or is it the balance of these two groups? Or is it diversity per se across all groups comprising a specific domain that is crucial?

In order to provide badly needed answers to these questions, this paper will explore the degree to which a wide variety of 1995 neighbourhood conditions in Sweden are statistically related to earnings for metropolitan and non-metropolitan men and women during the 1996–99 period, controlling for a wide variety of personal characteristics.

Alternative Mechanisms of Neighbourhood Effects and their Implications for what Household Mix Matters

There have been several comprehensive reviews of the potential theoretical links between neighbourhood processes and individual outcomes (in particular, see Atkinson *et al.*, 2001; Duncan *et al.*, 1997; Ellen & Turner, 2003; Friedrichs, 1998; Gephart, 1997; Haurin *et al.*, 2002; Jencks & Mayer, 1990; Sampson *et al.*, 2002). However, these theories often offer ambiguous or contradictory implications about what mix matters in neighbourhoods. The paper will describe these mechanisms only briefly here because they are well known,

drawing out any implications for household mix. In organizing this summary the useful distinction introduced by Manski (1995, 2000) among endogenous, exogenous, and correlated neighbourhood effects is employed.

Endogenous Neighbourhood Effects

Endogenous neighbourhood effects are those that occur when the behaviours or attitudes of one neighbourhood resident have a direct influence on (at least a portion of) his or her neighbours' behaviours or attitudes. Numerous versions of endogenous effects have been forwarded:

- *Socialization:* Behaviours and attitudes of all individuals may be changed (for better or worse) by contact with role models or peers who may be neighbours. When these changes occur they are often referred to as 'contagion effects'. For example, the actions by some to informally police and clean common neighbourhood spaces may encourage all others in the area to do the same.
- *Epidemic/Social Norms:* This is a special subset of socialization effects that are characterized by a minimum threshold being achieved before noticeable consequences ensue. The need for some subset of the neighbourhood population to reach a critical mass before their social norms begin to influence others to conform is a case in point. Another is the influence of local acts of crime and violence: when neighbours finally perceive the neighbourhood as too dangerous they will restrict their activities outside the home.
- *Selective Socialization:* This process is another special type of socialization process wherein neighbours are not all equally affected by others. Employed residents are often viewed as positive role models encouraging (only) their unemployed neighbours to find work, for example. Conversely, secondary school dropouts may discourage only their same-age peers from attending school and have no impact on other residents.
- *Social Networks:* Although it may be said that socialization proceeds through social networks, this is specified as a distinct process involving the communication of information and resources. One set of residents may intensify the density and multi-nodal structure of their social networks (create 'strong ties') by spatial clustering, thereby increasing their sources of assistance in times of need. On the other hand, such situations may lack the 'weak ties' that offer the prospect of bringing new information and resources into the community, thereby increasing social isolation.
- *Competition:* Under the premise that certain local resources are limited and not pure public goods, this theory posits that groups within the neighbourhood will compete for these finite resources amongst themselves. Because the context is a zero-sum game, social conflict will arise as one group more successfully competes. The control of a local public park for one type of specialized group activities provides an example.
- *Relative Deprivation:* This mechanism suggests that residents who have achieved some socio-economic success will be a source of disamenities for their less-well off neighbours. The latter will view the successful with envy or will make them perceive their own relative inferiority as a source of dissatisfaction.

What do these endogenous mechanisms imply about which household mix in neighbourhoods matters? The socialization, epidemic/social norms and selective socialization theories generally imply that neighbourhoods with larger shares of residents who exhibit 'socially undesirable' behaviours will encourage less favourable outcomes and those with larger shares of residents who exhibit 'socially desirable' behaviours will encourage more favourable outcomes for (some or all of) their neighbours. But this tells us little about which characteristics of the populace—demographic, socio-economic, ethnic, housing tenure—serve as the best proxies for these behaviours or what sort of mixing of groups is appropriate. The social networks theory is even more ambiguous in its implications, torn as it is between the potential benefits of strong ties fostered by own-group homogeneity versus those of extra-group weak ties, whose balance varies depending on the circumstance (Granovetter, 2005; Ionnides & Datcher Loury, 2004). The competition and relative deprivation theories are perhaps the most clear, suggesting that to mix socio-economically advantaged and disadvantaged residents in the same neighbourhood will further harm the latter group.

Exogenous Neighbourhood Effects

Exogenous neighbourhood effects occur if the behaviours or attitudes of one neighbour depend on the exogenous (or predetermined, fixed) characteristics of the individual's neighbours, such as ethnicity, religion or race. Here, for example, a recent immigrant may feel a special comfort and security because of proximity to another from the same national background, what is often termed 'ethnic solidarity'. As an illustration, Murie & Musterd (2004) note:

> Social homogeneity of a peripheral estate may offer an (temporary) asset, instead of a burden. Homogeneity may help to stimulate interaction and mutual support, which is especially important for newly arrived immigrants, who still have to learn their way in the place of settlement. (p. 1454)

Or, expressed in a less positive version, a person may have an aversion to proximity to certain prospective neighbours because of racial or religious differences and may therefore behave differently if confronted by them in the neighbourhood context. Yet another version of this mechanism may be termed 'social cohesion': the notion that residential contact among groups that differ in their exogenous characteristics will increase their social interactions and thereby reduce inter-group prejudices and misapprehensions, providing benefits not only to themselves but also to the larger society beyond the neighbourhood.

For analytical purposes here, exogenous neighbourhood effects are similar to the endogenous mechanisms described above, inasmuch as in both cases the characteristics of the households in the neighbourhood affect individual outcomes. Unfortunately, the exogenous theories are equally ambiguous in providing guidance as to what mix matters and in what direction.

Correlated Neighbourhood Effects

Correlated neighbourhood effect mechanisms do not vary by alterations in neighbourhood household composition, but rather are determined by larger structural forces in the

metropolitan area, like locations of jobs and geographic disamenities and the structures of local government. These external forces may impinge differentially on different neighbourhoods, but within any given neighbourhood they affect all residents roughly equally, producing thereby correlations in neighbours' outcomes. Several such mechanisms have been forwarded in the literature:

- Spatial mismatch: certain neighbourhoods have little accessibility (in either spatial proximity or as mediated by transportation networks) to job opportunities appropriate to the skills of their residents.
- Local institutional resources: certain neighbourhoods have access to few private, non-profit, or public institutions and organizations.
- Public services: certain neighbourhoods are located within local political jurisdictions that offer inferior services and facilities.
- External stigma: certain neighbourhoods may be stigmatized regardless of their current population because of their history, environmental or topographical disamenities, style, scale and type of dwellings, or condition of their commercial districts and public spaces.

These correlated effect mechanisms may be important,[3] but they will not be the subject of this paper. Rather, the focus is on the issue of neighbourhood population mix, and therefore on the endogenous and exogenous effect mechanisms arising from the household composition of the neighbourhood. It might be hypothesized that neighbourhood effects in a welfare state of the Swedish kind would be less pronounced compared to many other countries. This can be true but there are at least three reasons for the choice of Sweden here. First, the country has a legacy of anti-segregation policy within the framework of a redistributive welfare state. Such institutional arrangements may reduce the correlated effects but are they also effective in relation to endogenous and exogenous effects? Second, the segregation issue is high up the political agenda in Sweden, and the issue of social and ethnic mix is regarded as a crucial vehicle for avoiding the development of divided cities (Andersson, 2006). Finally, Sweden offers unique types of data, i.e. comprehensive longitudinal and geocoded information on all residents.

Previous Research on Neighbourhood Household Composition and Impacts on Individuals

Previous empirical research offers scant bases for answering the question of what mix matters. Although there have been numerous multivariate statistical studies attempting to quantify the relationship between indicators of neighbourhood household composition and a variety of outcomes for children, youth and adults (as reviewed in the articles cited above), relatively few have experimented with alternative measures of neighbourhood household composition in ways that permit meaningful comparisons of magnitudes and statistical significance of observed relationships. The results have been inconsistent, with some contingencies based on gender and race.

Most work has examined educational attainment. Corcoran *et al.* (1987) and Duncan (1994) found that measures of neighbourhood disadvantage (percentages of male unemployment, in poverty or using public assistance) were the strongest neighbourhood predictors, although the measures' strength depended on gender. Haveman & Wolfe (1994)

and Ginther *et al.* (2000) found that neighbourhood rates of not completing secondary school were the strongest correlates of an individual not completing secondary school. However, contrary results were produced by Crane (1991), Brooks-Gunn *et al.* (1993) and Ensminger *et al.* (1996), that emphasized the role played by affluent neighbours, at least for many gender-race categories.

Other studies have focused on teen sexual activity as the main outcome, generally finding that neighbourhood affluence was more predictive than neighbourhood disadvantage. Brewster's (1994) study determined that the share of advantaged neighbours was positively related to girls' use of contraception, but not boys'. Crane (1991) and Brooks-Gunn *et al.* (1993) found that a larger share of affluent neighbours reduced the likelihood of female teens' risk of bearing a child, but differed on whether the relationships were stronger for white or black female teens. However, the findings of South & Crowder (1999) were the opposite: the share of disadvantaged neighbours had a stronger relationship with teen childbearing outside marriage, as least for black women.

Finally in the area of labour market outcomes, Weinberg *et al.* (2004) discovered that several alternative measures of neighbours' social status were related to individuals' annual hours of work. In descending order of strength, these measures were rates of: public assistance, male employment, female employment, poverty and dropping out of secondary school. The impacts of neighbours' social status proved stronger for less-educated and Hispanic youth, although not for black (compared to white) youth. Dawkins *et al.* (2005) found that the percentage of poor neighbours was more predictive for individuals' duration of unemployment than either the percentage of secondary school dropouts or percentage of whites, although the coefficient was only statistically significant for blacks, not whites. However, they did not test the impact of affluent neighbours. Hoynes (2000) reported that the percentage of women in the postal code area who have never been married is a stronger predictor of individual women's entrances into and exits from public assistance in California than median household incomes. Van der Klaauw & Van Ours' (2003) study found native Dutch transition rates from welfare into work in Rotterdam were predicted by unemployment rates in the neighbourhood, but not a wide variety of other neighbourhood characteristics.

The extant work thus leaves a circumscribed and cloudy picture. Although most studies conclude that an individual's educational attainments and teenage fertility are best explained by the percentage of the 'advantaged' population in the neighbourhood (alternatively measured by professional-managerial occupations or high income), several find that characteristics of the 'disadvantaged' population (those on public assistance or not completing high school) are more important. However, labour market outcomes (annual hours of work, duration of unemployment), seem better explained by measures of 'disadvantaged' neighbours. Most studies stress the contingency of impacts based on race and gender. Moreover, to the authors' knowledge a comparison of alternative neighbourhood indicators has not been carried out in the realm of labour market earnings. Finally, the evidence thus far has been almost exclusively American. The study reported here hopes to contribute to filling this large gap in the literature by examining the statistical relationship between a large variety of neighbourhood characteristics and the subsequent earnings of Swedish men and women in both metropolitan and non-metropolitan settings, controlling for personal attributes and labour market characteristics.

Empirical Approach

Data Sources

Most of the variables employed were 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, labour market and population). The current study merged selected information about individuals from annual *Louise* files to create a longitudinal database 1991–99 for all individuals present in Sweden in 1991. The other sources of information used were the Statistics Sweden real estate and property registers, from which the study obtained data on housing type and tenure that were used to construct certain neighbourhood characteristic variables.

It must be emphasized that the dataset includes observations of virtually the *entire population* within the desired adult age range, not a sample. Thus, the *t*-statistics presented below should not be interpreted as guides for prospective errors involving inferences from a sample to the larger population.

Model for Conducting Empirical Tests

The outcome of interest here is the average annual income from work (during 1996–99) for working-age individuals (ages 20–60 in 1996) who were residents of Sweden in 1991 and present each year 1995–99.[4] Since this indicator encapsulates educational credentials, labour force participation, employment regularity and hourly compensation, it is thought to be the most comprehensive single measure of an individual's economic worth. Descriptive statistics for this outcome variable are presented for males in Table 1 and for females in Table 2, both stratified by metropolitan/non-metropolitan place of residence in 1995.

The study models income as a gender- and place-specific, log-linear function of personal characteristics, characteristics of the neighbourhood in which they reside at the beginning of the period, municipality of residence at the beginning of the period, and local labour market conditions at the end of the period in question.[5] Symbolically:

$$\ln(I_{96-99ij}) = \alpha + \beta[P_{96-99ij}] + \gamma[P_{95ij}] + \theta N_{95ij} + \mu[M_{95ij}] + \phi[L_{95-99ij}] + \epsilon \quad (1)$$

where:

I_{96-99} = cumulative income from work (annualized) observed for individual during period 1996–99.[6]

$[P_{96-99}]$ = observed personal characteristics that can vary over time (e.g. marital or fertility status, educational attainment).

$[P_{95}]$ = observed personal characteristics that do not vary after 1995 (e.g. year and country of birth, experiences prior to 1995).

N_{95} = observed characteristics of neighbourhood where individual resides in 1995 (a variety of characteristics will be modelled).

$[M_{95}]$ = municipality of residence in 1995.

$[L_{95-99}]$ = observed characteristics of local labour market(s) in which individual resides at 1995 and 1999 (e.g. mean earnings).

ϵ = a random error term with the usual assumed statistical properties

i = gender

j = metropolitan or non-metropolitan place

Table 1. Descriptive statistics for outcome and control variables[a]: males

	Metro residence		Non-Metro residence	
	Mean	Std. dev.	Mean	Std. dev.
Outcome Variable				
Annual mean labour income, 1996–99 (SEK100)	2175.54	1839.08	1893.63	1222.39
Control Variables				
No. children under age 7, 1995	0.31	0.66	0.32	0.68
Some sick leave during 1995 (1 = yes)	0.13	0.34	0.15	0.35
Pre-retired during 1995 (1 = yes)	0.05	0.22	0.06	0.23
Parental leave during 1995 (1 = yes)	0.17	0.38	0.19	0.39
Studying during 1995 (1 = yes)	0.05	0.22	0.03	0.17
No. years with pre-retirement, 1996–99 (1 = yes)	0.27	0.96	0.29	1.01
No. child-years under 7, 1996–99	1.12	2.28	1.09	2.27
No. years studying, 1996–99	0.16	0.62	0.12	0.55
No. years with parental leave, 1996–99	0.63	1.23	0.64	1.24
No. years with sick leave, 1996–99	0.44	0.95	0.53	1.04
Immigrants w/ < 5 years in Sweden (1 = yes)	0.008	0.089	0.004	0.061
Western country of birth (1 = yes)	0.07	0.25	0.04	0.21
Eastern European country of birth (1 = yes)	0.03	0.16	0.01	0.11
Non-Western country of birth (1 = yes)	0.06	0.24	0.02	0.14
No formal education (1 = yes)	0.01	0.11	0.01	0.07
< 10 years education (1 = yes)	0.09	0.29	0.16	0.36
10 years education (1 = yes)	0.13	0.34	0.14	0.34
13 years, some post-secondary (1 = yes)	0.10	0.30	0.07	0.25
14+ years, but no PhD (1 = yes)	0.20	0.40	0.13	0.34
PhD attained (1 = yes)	0.017	0.130	0.004	0.066
Education rose LT 11–12 to 11–12 + (1 = yes)	0.005	0.071	0.003	0.054
Education rose 11–12 to higher (1 = yes)	0.01	0.10	0.01	0.10
Age in years	40.96	10.28	41.64	10.33
Age more than 50 years (1 = yes)	0.22	0.42	0.24	0.43
Single 1995 but couple 1999 (1 = yes)	0.08	0.27	0.06	0.24
Couple 1995 but single 1999 (1 = yes)	0.06	0.24	0.06	0.24
Single 1991 but couple 1995 (1 = yes)	0.09	0.29	0.08	0.27
Couple 1991 but single 1995 (1 = yes)	0.09	0.28	0.08	0.27
Mean income in local labour market, 1999	1769.72	144.21	1550.79	90.16
N	847162		1236095	

Note: [a] The listed variables were used in the model runs. Omitted variables, e.g. 11–12 years of education, are reference categories. For Metro males, 45 per cent belong to the educational reference category. For Non-Metro males the corresponding value is 49 per cent.

All Greek letters represent parameters to be estimated through OLS multiple regression techniques, with each regression stratified for a particular gender/place stratum.

The study will compare the magnitudes of the (standardized) coefficients θ and their statistical significance across various specifications of N within each of four strata as the empirical test. All other control variables in (1) shall remain the same in these trials.

Table 2. Descriptive statistics for outcome and control variables[a]: females

	Metro residence		Non-Metro residence	
	Mean	Std. dev.	Mean	Std. dev.
Outcome Variable				
Annual mean labour income, 1996–99 (SEK100)	1520.57	1024.81	1293.88	819.39
Control Variables				
No. children under age 7, 1995	0.35	0.69	0.37	0.70
Some sick leave during 1995 (1 = yes)	0.19	0.39	0.20	0.40
Pre-retired during 1995 (1 = yes)	0.07	0.25	0.09	0.28
Parental leave during 1995 (1 = yes)	0.28	0.45	0.29	0.45
Studying during 1995 (1 = yes)	0.07	0.26	0.05	0.22
No. years with pre-retirement, 1996–99 (1 = yes)	0.35	1.09	0.43	1.20
No. child-years under 7, 1996–99	1.23	2.36	1.18	2.34
No. years studying, 1996–99	0.28	0.79	0.27	0.78
No. years with parental leave, 1996–99	0.98	1.52	0.96	1.51
No. years with sick leave, 1996–99	0.66	1.07	0.75	1.14
Immigrants w/ < 5 years in Sweden (1 = yes)	0.008	0.086	0.004	0.061
Western country of birth (1 = yes)	0.08	0.27	0.05	0.22
Eastern European country of birth (1 = yes)	0.03	0.18	0.01	0.12
Non-Western country of birth (1 = yes)	0.05	0.21	0.02	0.13
No formal education (1 = yes)	0.007	0.086	0.003	0.058
< 10 years education (1 = yes)	0.08	0.27	0.12	0.33
10 years education (1 = yes)	0.12	0.32	0.12	0.33
13 years, some post-secondary (1 = yes)	0.06	0.24	0.03	0.17
14+ years, but no PhD (1 = yes)	0.28	0.45	0.21	0.41
PhD attained (1 = yes)	0.005	0.074	0.001	0.034
Education rose LT 11–12 to 11–12 + (1 = yes)	0.004	0.065	0.003	0.057
Education rose 11–12 to higher (1 = yes)	0.02	0.13	0.02	0.13
Age in years	41.01	10.33	41.95	10.32
Age more than 50 years (1 = yes)	0.22	0.42	0.25	0.43
Single 1995 but couple 1999 (1 = yes)	0.07	0.25	0.05	0.22
Couple 1995 but single 1999 (1 = yes)	0.06	0.24	0.06	0.24
Single 1991 but couple 1995 (1 = yes)	0.08	0.28	0.08	0.26
Couple 1991 but single 1995 (1 = yes)	0.08	0.27	0.07	0.25
Mean income in local labour market, 1999	1770.51	144.18	1550.22	87.74
N	842542		1181080	

Note: [a] The listed variables were used in the model runs. Omitted variables, e.g. 11–12 years of education, are reference categories. For Metro females, 45 per cent belong to the educational reference category. For Non-Metro females the corresponding value is 52 per cent.

Variables

Measures of neighbourhood composition. In this study 'neighbourhood' is operationalized 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. There are approximately 9000 SAMS

in Sweden, each with an average population of approximately 1000 inhabitants; SAMS in metropolitan areas are considerably smaller geographically than those in non-metropolitan areas, prompting the study to stratify the analyses by place so that geographic scale of neighbourhood can be made more comparable across individuals being analyzed.[7]

It is important to test for three aspects of neighbourhood composition: (1) 'absolute' share of a 'particular' group, (2) 'relative' shares of two 'particular' groups; and (3) 'overall' diversity among 'all' groups. The study measures the: (1) absolute share as percentages of the 'advantaged' and 'disadvantaged' categories; (2) the relative share as the ratio of the percentages of these two categories; and (3) overall diversity as an entropy index based on all categories of data available. The entropy index measures the extent to which a population is *evenly* distributed among the specified categories (Theil, 1972) and is calculated as:

$$E = \sum_{m=1}^{M} \pi_m \ln\left(\frac{1}{\pi_m}\right),$$

where π_m denotes the proportion of individuals in category m in the neighbourhood. The entropy index ranges between 0 and ln M: 0 when all individuals are members of a single category; ln M when individuals are evenly distributed among the M categories. E is divided by its maximum value ln M to set the upper bound to 1. The theoretical and empirical superiority of this measure of inequality has been demonstrated by Reardon & Firebaugh (2002).

These three versions of composition measures are computed in four domains for each SAMS in Sweden in 1995:

- *Education:* (1) percentage aged 20 years and older with 'low' (less than 12 years of schooling) or 'high' (14 or more years) educational attainment; (2) the ratio of these; (3) entropy based on seven educational attainment groups (see Appendix Table 1A).
- *Ethnicity:* (1) percentage of foreign-born and percentage native-born Swedes; (2) the ratio of these; (3) entropy based on four regional places of birth: Swedes, Western, East European, other/non-Western.
- *Income:* (1) percentage lowest income males (0–29th percentiles), percentage high-income males (70th percentile or higher); (2) ratio of these; (3) entropy based on all male income deciles.
- *Housing tenure:* (1) percentage households living in public rental housing,[8] percentage in owner-occupied units; (2) the ratio of these; (3) entropy based on four housing type categories: those above, plus cooperatives and private rental.

Sweden has pursued a social mix neighbourhood policy since the mid-1970s. Its efficiency may be questioned (Musterd & Andersson, 2005) and its rather vague goal formulation but there is no doubt that the primary instrument has been tenure mix and that the idea has been to avoid segregation by getting a mix in terms of households' social class positions and (later) also ethnic origin (Bergsten & Holmqvist, 2007). Hence, the choice here of the four dimensions is of policy relevance but it is also relevant from the perspective of existing theory concerning endogenous and exogenous neighbourhood effects.

In computing neighbourhood population characteristics, all 1995 residents of the SAMS age 20–65 years are used; in computing housing characteristics, all housing units in the SAMS in 1995 are used.

Control variables. The personal characteristics of individuals [$P_{96-99ij}$] and [P_{95ij}] are operationalized 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 probably not related to neighbourhood context (such as parental leave or pre-retirement status). [M_{95ij}] is operationalized with a set of dummy variables representing each municipality of residence in 1995 and [L_{99ij}] with the mean earnings for the local labour market area in which the individual resided in 1999. Descriptive statistics for these control variables are presented for males in Table 1 and for females in Table 2, both stratified by metropolitan/non-metropolitan place of residence in 1995.

Results

Relationships among Neighbourhood Mix Measures

Although the aforementioned variables for neighbourhood household mix are distinct conceptually, they may prove not to be in practice. If it proved the case that all measures were highly correlated the question of what mix matters would be rendered moot. This clearly is not the case here, as shown by the bivariate correlation coefficients in Table 3. As expected, some are of course highly correlated, in particular those related to the ethnic composition; high proportions of natives means of course low presence of immigrants, and vice versa. However, in general correlations are low or modest.

Patterns of Neighbourhood Mix in Sweden

Neighbourhood characteristics are displayed for all groups separately in Table 4 (metropolitan) and Table 5 (non-metropolitan). The metropolitan subset contains all neighbourhoods in the Stockholm, Gothenburg and Malmö labour market regions, comprising 35 per cent of the SAMS areas and about 41 per cent of the population in the selected age groups (aged 20–64 in 1995). Hence, Table 5 shows the neighbourhood characteristics for the rest of Sweden. Values for the mean, standard deviation, 1st percentile, and 99th percentile are presented for the four subsets.

Entropy values for incomes and education are very compressed (mean approximately 0.9), indicating highly diverse neighbourhoods in the country in terms of class compositions (Musterd & Andersson, 2005). This is to be expected, given the long-standing Swedish policy of building developments with a mix of housing types and providing social housing benefits to a large share of the households. Also as expected, values for ethnic and housing tenure entropies show wider ranges (means 0.25–0.40 and around 0.51 respectively), with lower means for the ethnic entropies in non-metropolitan areas.

The range of values is much greater for other measures. Immigrants are more often found in metropolitan regions, resulting both in higher averages and higher values for the 99th percentile in metropolitan regions. Also as expected, the top values for percentage high-income males and percentage highly educated are found in metropolitan regions. Homeownership is more common in non-metropolitan areas (60/40), but entropy measures as well as percentage values for tenure forms show substantial variation across neighbourhoods in both types of regional settings.

Table 3. Correlations between the neighbourhood characteristics: Swedish males

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	Ethnic entropy		0.76	−0.95	0.67	−0.25	0.49	−0.21	0.11	0.02	0.23	−0.11	0.17	0.31	0.53	−0.58	0.33
2	% non-Western 1995	0.79		−0.79	0.95	−0.25	0.57	−0.29	0.18	0.02	0.13	−0.01	0.09	0.24	0.53	−0.56	0.43
3	% Swedes 1995	−0.92	−0.93		−0.76	0.30	−0.55	0.25	−0.15	0.03	−0.25	0.13	−0.20	−0.24	−0.50	0.53	−0.35
4	Ratio % non-Western / % Swedes	0.53	0.86	−0.78		−0.29	0.53	−0.28	0.18	−0.02	0.17	−0.05	0.14	0.15	0.44	−0.45	0.36
5	Income entropy	−0.03	−0.17	0.18	−0.30		0.03	−0.43	−0.22	0.10	0.12	−0.26	0.02	0.22	−0.09	−0.07	−0.17
6	% in lowest 3 male income deciles	0.67	0.72	−0.75	0.59	−0.02		−0.69	0.36	−0.21	0.21	−0.14	0.23	0.28	0.44	−0.55	0.39
7	% in highest 3 male income deciles	−0.55	−0.54	0.58	−0.41	−0.44	−0.81		−0.23	0.28	−0.52	0.52	−0.46	−0.26	−0.28	0.36	−0.24
8	Ratio % lowest 3 / % highest 3 income deciles	0.34	0.46	−0.50	0.61	−0.33	0.56	−0.40		−0.21	0.03	0.05	0.07	0.04	0.17	−0.14	0.17
9	Educational entropy	0.10	0.05	−0.04	0.00	0.32	−0.14	−0.07	−0.10		−0.24	0.40	−0.59	0.01	−0.08	0.02	−0.08
10	% low educated	0.46	0.52	−0.54	0.47	0.17	0.41	−0.64	0.31	0.28		−0.85	0.85	−0.09	0.08	0.11	0.04
11	% high educated	−0.28	−0.26	0.29	−0.22	−0.31	−0.11	0.47	−0.12	−0.43	−0.88		−0.77	0.07	−0.02	−0.12	0.03
12	Ratio % low education / % high education	0.41	0.46	−0.49	0.43	0.07	0.40	−0.57	0.31	−0.05	0.90	−0.76		−0.07	0.08	0.07	0.05
13	House type entropy	0.13	0.04	−0.06	−0.01	0.34	0.16	−0.22	0.02	0.04	−0.10	0.13	−0.11		0.26	−0.69	0.05
14	% public rental	0.60	0.58	−0.57	0.39	0.10	0.53	−0.52	0.23	0.09	0.38	−0.23	0.32	0.04		−0.56	0.56
15	% owner occupied	−0.61	−0.44	0.52	−0.26	−0.34	−0.62	0.63	−0.22	0.02	−0.12	0.00	−0.13	−0.52	−0.51		−0.34
16	Ratio % public rental / % owner occupied	0.35	0.47	−0.42	0.44	0.02	0.35	−0.28	0.16	0.03	0.18	−0.04	0.15	0.12	0.47	−0.35	

Note: The coefficients below the diagonal are for metropolitan males and the coefficients above the diagonal are for non-metropolitan males.

Table 4. Descriptive statistics for neighbourhood variables in metropolitan areas

			Percentiles	
	Mean	Std. dev.	1	99
Males				
Ethnic entropy	0.405	0.186	0.110	0.885
% non-Western 1995	0.062	0.095	0.000	0.525
% Swedes 1995	0.859	0.110	0.415	0.974
Ratio % non-Western / % Swedes	0.127	0.388	0.000	1.981
Income entropy	0.900	0.068	0.669	0.975
% in lowest 3 male income deciles	0.307	0.129	0.140	0.762
% in highest 3 male income deciles	0.345	0.141	0.050	0.636
Ratio % lowest 3 / % highest 3	1.574	4.314	0.242	13.519
Educational entropy	0.910	0.053	0.692	0.979
% low educated	0.231	0.096	0.062	0.503
% high educated	0.312	0.145	0.093	0.697
Ratio % low / % high	1.107	1.048	0.087	4.953
House type entropy	0.518	0.291	0.000	0.986
% public rental	0.187	0.279	0.000	1.000
% owner occupied	0.398	0.392	0.000	1.000
Ratio % public rental / % owner occupied	213.6	613.7	0.0003	3465.0
Females				
Ethnic entropy	0.403	0.183	0.114	0.879
% non-Western 1995	0.061	0.091	0.000	0.525
% Swedes 1995	0.861	0.105	0.432	0.973
Ratio % non-Western / % Swedes	0.119	0.360	0.000	1.981
Income entropy	0.900	0.067	0.684	0.975
% in lowest 3 male income deciles	0.300	0.120	0.138	0.710
% in highest 3 male income deciles	0.352	0.139	0.055	0.638
Ratio % lowest 3 / % highest 3	1.452	3.839	0.242	12.337
Educational entropy	0.911	0.052	0.704	0.978
% low educated	0.228	0.094	0.062	0.503
% high educated	0.315	0.144	0.094	0.692
Ratio % low / % high	1.076	1.018	0.088	4.799
House type entropy	0.520	0.292	0.000	0.986
% public rental	0.186	0.278	0.000	1.000
% owner occupied	0.400	0.391	0.000	0.978
Ratio % public rental / % owner occupied	211.0	601.1	0.0003	2572.5

Multivariate Regression Results

Table 6 shows the results of the regression analyses for the neighbourhood variables in metropolitan areas and Table 7 shows the results in non-metropolitan areas. The conclusions to be made here are based mainly on the standardized regression coefficients (Std. Beta) that show how many standard deviations (along the natural log scale) the dependent variable changes when the neighbourhood variable increases by one standard deviation. Statistical significance of the effects is not shown, because given the size of the dataset, the p value for all effects is below 0.001, except for educational entropy

Table 5. Descriptive statistics for neighbourhood variables in non-metropolitan areas

			Percentiles	
	Mean	Std. dev.	1	99
Males				
Ethnic entropy	0.257	0.150	0.051	0.746
% non-Western 1995	0.022	0.038	0.000	0.209
% Swedes 1995	0.923	0.066	0.646	0.987
Ratio % non-Western / % Swedes	0.029	0.069	0.000	0.330
Income entropy	0.933	0.042	0.782	0.982
% in lowest 3 male income deciles	0.284	0.099	0.122	0.621
% in highest 3 male income deciles	0.277	0.107	0.061	0.558
Ratio % lowest 3 / % highest 3	1.500	5.391	0.261	8.691
Educational entropy	0.905	0.038	0.799	0.971
% low educated	0.285	0.084	0.108	0.475
% high educated	0.207	0.093	0.072	0.501
Ratio % low / % high	1.819	1.299	0.220	5.867
House type entropy	0.511	0.224	0.000	0.961
% public rental	0.107	0.181	0.000	0.886
% owner occupied	0.593	0.311	0.000	0.998
Ratio % public rental / % owner occupied	21.4	121.4	0.0004	702.0
Females				
Ethnic entropy	0.258	0.149	0.052	0.746
% non-Western 1995	0.023	0.037	0.000	0.199
% Swedes 1995	0.923	0.065	0.653	0.987
Ratio % non-Western / % Swedes	0.029	0.067	0.000	0.328
Income entropy	0.933	0.041	0.783	0.982
% in lowest 3 male income deciles	0.281	0.094	0.122	0.594
% in highest 3 male income deciles	0.281	0.107	0.068	0.564
Ratio % lowest 3 / % highest 3	1.390	3.314	0.251	7.395
Educational entropy	0.906	0.037	0.802	0.972
% low educated	0.283	0.084	0.109	0.471
% high educated	0.210	0.094	0.074	0.500
Ratio % low / % high	1.777	1.268	0.225	5.722
House type entropy	0.512	0.225	0.000	0.961
% public rental	0.108	0.181	0.000	0.886
% owner occupied	0.594	0.309	0.000	0.999
Ratio % public rental / % owner occupied	21.0	119.6	0.0004	702.0

for metropolitan males ($p = 0.579$). The common list of control variables used in all models can be seen in Appendix Table 1A, which also shows the coefficients for these variables in one representative model (effect of the proportion in lowest three male income deciles for females in non-metropolitan areas).[9]

If the strength of neighbourhood effects is compared across the four strata, it can be seen that the effects are generally stronger in metropolitan than in non-metropolitan areas and stronger for males than for females, regardless of indicator. However, it is specifically the extremes of the income distribution in the neighbourhood that matter most, while the mix

Table 6. Effects of neighbourhood characteristics[a] in metropolitan areas when added to a model controlling for individual, municipality and labour market characteristics

	B	Std. Beta	Exp(B)	R squared
Males				
Ethnic entropy	−0.8062	−0.067	0.45	0.456
% non-Western 1995	−1.5451	−0.065	0.21	0.456
% Swedes 1995	1.5548	0.076	4.73	0.457
Ratio % non-Western / % Swedes	−0.3044	−0.053	0.74	0.455
Income entropy	−0.1557	−0.005	0.86	0.453
% in lowest 3 male income deciles	−2.1475	−0.124	0.12	0.466
% in highest 3 male income deciles	1.6957	0.107	5.45	0.463
Ratio % lowest 3 / % highest 3	−0.0305	−0.059	0.97	0.456
Educational entropy	−0.0206	0.000	0.98	0.453
% low educated	−1.4476	−0.062	0.24	0.456
% high educated	0.6047	0.039	1.83	0.454
Ratio % low / % high	−0.1159	−0.054	0.89	0.455
House type entropy	−0.1254	−0.016	0.88	0.455
% public rental	−0.3818	−0.048	0.68	0.457
% owner occupied	0.4144	0.073	1.51	0.459
Ratio % public rental / % owner occupied	−0.0001	−0.032	1.00	0.454
Females				
Ethnic entropy	−0.5572	−0.048	0.57	0.452
% non-Western 1995	−1.3602	−0.057	0.26	0.453
% Swedes 1995	1.2787	0.063	3.59	0.454
Ratio % non-Western / % Swedes	−0.3175	−0.053	0.73	0.453
Income entropy	0.4758	0.015	1.61	0.451
% in lowest 3 male income deciles	−1.4066	−0.078	0.24	0.456
% in highest 3 male income deciles	0.8997	0.058	2.46	0.454
Ratio % lowest 3 / % highest 3	−0.0273	−0.049	0.97	0.453
Educational entropy	0.2456	0.006	1.28	0.451
% low educated	−1.1583	−0.051	0.31	0.453
% high educated	0.3642	0.025	1.44	0.451
Ratio % low / % high	−0.1020	−0.048	0.90	0.452
House type entropy	−0.0264	−0.004	0.97	0.452
% public rental	−0.2774	−0.036	0.76	0.453
% owner occupied	0.1773	0.032	1.19	0.452
Ratio % public rental / % owner occupied	−0.0001	−0.029	1.00	0.452

Note: [a]The 16 neighbourhood variables are in the model with the control variables one at a time. This Table is thus the result of 32 runs where only the neighbourhood variable is exchanged.

in the domains of housing tenure, ethnicity and education have progressively weaker effects. Indeed, variables measuring the share of both low-income and high-income males in the neighbourhood have the strongest effects in both metropolitan and non-metropolitan areas and for both sexes. The standardized beta coefficients for percentage low-income range from − 0.124 for metropolitan males to − 0.054 for non-metropolitan females; the comparable figures for percentage high-income are 0.107 and 0.053.

When the effects of different mix measures of each domain are compared, it can be seen that the proportions of specific groups have stronger effects than either general diversity

Table 7. Effects of neighbourhood characteristics[a] in non-metropolitan areas when added to a model controlling for individual, municipality and labour market characteristics

	B	Std. Beta	Exp(B)	R squared
Males				
Ethnic entropy	−0.6102	−0.044	0.54	0.460
% non-Western 1995	−2.5430	−0.046	0.08	0.460
% Swedes 1995	1.5245	0.048	4.59	0.460
Ratio % non-Western / % Swedes	−1.2747	−0.042	0.28	0.460
Income entropy	−0.6216	−0.013	0.54	0.459
% in lowest 3 male income deciles	−1.9630	−0.093	0.14	0.466
% in highest 3 male income deciles	1.7606	0.090	5.82	0.465
Ratio % lowest 3 / % highest 3	−0.0051	−0.013	0.99	0.459
Educational entropy	1.0698	0.019	2.91	0.459
% low educated	−1.4961	−0.060	0.22	0.460
% high educated	1.0794	0.048	2.94	0.460
Ratio % low / % high	−0.0717	−0.045	0.93	0.460
House type entropy	−0.1713	−0.018	0.84	0.459
% public rental	−0.3381	−0.029	0.71	0.460
% owner occupied	0.3019	0.045	1.35	0.460
Ratio % public rental / % owner occupied	−0.0004	−0.026	1.00	0.459
Females				
Ethnic entropy	−0.4229	−0.029	0.66	0.458
% non-Western 1995	−1.9960	−0.034	0.14	0.458
% Swedes 1995	1.1756	0.035	3.24	0.458
Ratio % non-Western / % Swedes		−0.034	0.34	0.459
Income entropy	−0.1590	−0.003	0.85	0.458
% in lowest 3 male income deciles	−1.2448	−0.054	0.29	0.460
% in highest 3 male income deciles	1.0663	0.053	2.90	0.460
Ratio % lowest 3 / % highest 3	−0.0111	−0.017	0.99	0.458
Educational entropy	1.0198	0.018	2.77	0.458
% low educated	−1.3660	−0.053	0.26	0.459
% high educated	0.8872	0.038	2.43	0.458
Ratio % low / % high	−0.0704	−0.041	0.93	0.459
House type entropy	−0.0684	−0.007	0.93	0.458
% public rental	−0.2293	−0.019	0.80	0.458
% owner occupied	0.1319	0.019	1.14	0.458
Ratio % public rental / % owner occupied	−0.0004	−0.020	1.00	0.458

Note: [a]The 16 neighbourhood variables are in the model with the control variables one at a time. This Table is thus the result of 32 runs where only the neighbourhood variable is exchanged.

as measured by the entropy indices, or ratio of disadvantaged and advantaged groups (although this is less clear for education). Of the entropy measures, only ethnic entropy has an effect comparable to the effect of the proportion variable. However, in the Swedish context, high ethnic entropy actually means high proportion of immigrants, while low ethnic entropy can be observed mainly in neighbourhoods with mostly Swedish population. This means that the entropy measure may actually capture the effect of the proportion of immigrants in the neighbourhood. This may also explain why ethnic entropy has the strongest effect of the entropy measures.

Table 8. Predicted incomes in different neighbourhood conditions

Cut-point of the neighbourhood variable	% lowest 3 male income deciles at the cut-point	Predicted average labour income 1996–99 (SEK100)	% non-Western 1995 at the cut-point	Predicted average labour income 1996–99 (SEK100)
Males, *metropolitan*				
5th percentile	16.7	2855	0.3	2318
Mean − std. dev.	17.8	2788	<0	–
Mean	30.7	2114	6.2	2114
Mean + std. dev.	43.5	1603	15.7	1827
95th percentile	58.9	1152	23.7	1614
Females, *metropolitan*				
5th percentile	16.5	1902	0.3	1701
Mean − std. dev.	18.0	1862	<0	–
Mean	30.0	1573	6.1	1573
Mean + std. dev.	42.0	1329	15.1	1390
95th percentile	54.1	1120	22.9	1250
Males, *non-metropolitan*				
5th percentile	15.9	2562	0.0	2121
Mean − std. dev.	18.5	2433	<0	–
Mean	28.4	2004	2.2	2004
Mean + std. dev.	38.3	1650	6.0	1820
95th percentile	46.8	1396	8.1	1725
Females, *non-metropolitan*				
5th percentile	15.8	1605	0.0	1441
Mean − std. dev.	18.6	1550	<0	–
Mean	28.1	1378	2.3	1378
Mean + std. dev.	37.5	1225	5.9	1280
95th percentile	45.7	1106	8.1	1227

A further inspection of the effects of the proportions of specific groups reveals that it is mostly the proportion of the disadvantaged group rather than the proportion of the advantaged group that has the stronger effect, although the differences are not large. For the income and education variables the proportion of the disadvantaged group has the stronger effect in all strata. However, for the housing tenure and ethnic variables there is no clear pattern.

The magnitudes of estimated relationships between selected neighbourhood variables and individuals' subsequent labour earnings can be seen in more concrete terms in Table 8, which is based on the multivariate analyses. It shows the predicted labour income for an average individual when the proportion in three lowest male income deciles and the proportion of foreign born vary between their extremes. These predicted incomes are calculated for a hypothetical individual with mean characteristics, whose average yearly labour income in 1996–99 would be the median income in his/her stratum when living in an average neighbourhood in terms of proportion of low-income or non-Western neighbours. The ratios of predicted incomes between different neighbourhood conditions would be the same as in Table 8 for other types of individuals, too.

The main observation from Table 8 is that the differences in the predicted labour income between the extreme neighbourhood conditions (represented by the 5th and 95th

percentiles of the variables) are huge, especially in metropolitan areas. Because using these extremes might overemphasize the strength of the effects, the effects of one-standard-deviation deviations from the average values are also shown. The differences in the predicted labour incomes are large nevertheless, even when these cut-points are used. As an illustration, for the largest apparent effect of percentage of low-income in a neighbourhood (males in metropolitan areas), a difference is seen in the 1996–99 average labour income of SEK118 000 (56 per cent of the median) associated with 1995 residence in neighbourhoods that are one standard deviation above and below the mean, respectively. For the smallest apparent effect of percentage of low-income in a neighbourhood (females in non-metropolitan areas) the comparable figures are SEK32 000 (24 per cent of the median).

A final comment is offered to provide the context in interpreting these perhaps unexpectedly large differences in predicted incomes. It should be noted that, although large in magnitude, the neighbourhood variables do not add much to the proportion of explained variance in the models.[10] This means that although there are large differences between predicted incomes associated with neighbourhood income variations, this effect is far from deterministic. There is much variation around the predicted values depending on individual characteristics and other contextual factors (many of which are unmeasured), such as family background. This issue of omitted variables is returned to below.

Extensions of the Analysis

To further probe the issue of what mix matters, the basic model was extended to examine how two measures of neighbourhood composition would perform in the *same* regression. There was particular interest in exploring whether the similarity in coefficients for the percentage of low-income and percentage of high-income males in the neighbourhood would persist when they were both entered into the regression. It was found that they diverged substantially in importance for all subsets (Table 9). For all four strata, the effect of low-income residents proved to be larger than the effect of high-income residents in the neighbourhood. For metropolitan females, the effect of high-income males was slightly

Table 9. Effects of the proportions of low- and high-income residents when simultaneously in the model

	B	Std. Beta	Exp(B)	R squared
Males, *metropolitan*				
% in lowest 3 male income deciles	−1.8437	−0.106	0.16	0.466
% in highest 3 male income deciles	0.3428	0.022	1.41	
Females, *metropolitan*				
% in lowest 3 male income deciles	−1.6157	−0.090	0.20	0.456
% in highest 3 male income deciles	−0.2213	−0.014	0.80	
Males, *non-metropolitan*				
% in lowest 3 male income deciles	−1.3926	−0.066	0.25	0.466
% in highest 3 male income deciles	0.7167	0.037	2.05	
Females, *non-metropolitan*				
% in lowest 3 male income deciles	−0.8874	−0.039	0.41	0.460
% in highest 3 male income deciles	0.4263	0.021	1.53	

negative in this model, but for the other strata the effect of high-income males stayed positive and the effect of low-income males negative.

Econometric Issues

The data are organized across multiple spatial levels: the individual, the SAMS, the municipality and the local labour market. It is recognized that the use of OLS under such circumstances is inefficient and produces standard errors that are biased downward. Nevertheless, the *t*-statistics are incredibly robust and therefore the authors are confident that they are not drawing improper inferences by using them.

All models were also checked for multicollinearity, and it was found that none used standard benchmarks for tolerance and variance inflation factors (Belsey *et al.*, 1980).

Finally, the issue of potential selection bias is addressed. The coefficients of neighbourhood characteristics may not be fully appropriate estimates of causal effects insofar as unobserved individual characteristics that affect labour income may also affect their choice of neighbourhood (Galster, 2003). Although the direction of this potential bias cannot be ascertained with certainty, it is acknowledged that the point estimates cannot be interpreted as indicative of purely causal influences emanating from the neighbourhood environment. Nevertheless, it can be argued that this should not detract from the tests being performed in this paper. Whatever the unobserved selection bias, it should affect equally the coefficients of all the estimated neighbourhood mix variables. Therefore, conclusions regarding the rank ordering of which measure provides the greatest explanatory power should be unaffected by selection bias.

Conclusions and Caveats

It is recognized that neighbourhood household mix itself typically is not the desired end by policy makers, and that indicators of mix may be poor proxies for the neighbourhood social processes that we probably do care about (such as collective efficacy, bridging social capital networks, community cohesion, etc.). Although there are correlations between neighbourhood population mix and some underlying processes (Cook *et al.*, 1997; Sampson, 1997; Sampson *et al.*, 1999), this is not the point here. Rather, it is that policy makers (especially in countries like Sweden, Denmark and the Netherlands) are taken at their word—that they are aiming for mix as a proximal outcome. In this context it is perfectly appropriate to test the extent to which such proximal mix indicators in fact correlate with a widely accepted measure of economic success: labour income.

The study found that how household mix is operationalized does indeed matter. The extremes of the neighbourhood income distribution in 1995, operationalized by the percentages of adult males with earnings in the lowest 30th and the highest 30th percentiles, hold greater explanatory power for average labour income in 1996–99 than domains of household mix related to education, ethnicity or housing tenure. This holds for both genders residing in either metropolitan or non-metropolitan areas of Sweden. By contrast, diversity per se (as measured by entropy) has a much weaker relationship with economic outcomes.

Without doubt and in principle, the findings here support two types of policy interventions already on the agenda in some Western European countries. In Sweden and the Netherlands, for example, the presumed existence of neighbourhood effects at least

implicitly underpins urban area-based programmes launched to support residents in deprived neighbourhoods (Andersson & Musterd, 2005). Such programmes do have different points of departures and rationales but if it is proved that people residing in neighbourhoods having a high concentration of low-income people are indeed negatively affected by their residency as such, politicians can argue that extra resources have to be added in order to compensate for the negative effects. Furthermore, these two countries, but also the UK, Finland and others, have active housing and social mix programmes as an important part of their current housing policies. However, such programmes seldom make clear what mix is desirable and appropriate, only that mix is good. Therefore, clarifying what mix matters is seen here as an important task for social science research. According to research findings in a recently finished EU-funded project, Urban Governance, Inclusion and Sustainability (UGIS), both area-based policies and most mix policies are now partly driven by the fear of ethnic clustering (Andersson, 2003; Beaumont *et al.*, 2003). The findings of the current study do not support the hypothesis that the ethnic dimension is the most crucial one. On the contrary, the study finds that the socio-economic composition of neighbourhoods is the most important dimension, at least in terms of individuals' incomes.

It must be emphasized that the findings should not be misinterpreted as advocating for a particular sort of neighbourhood intervention strategy or what socio-economic mix of neighbourhoods is optimal. Such an analysis requires a much deeper consideration of the social equity and efficiency bases of the mixing policy (Galster, 2005, 2007a, 2007b), which is beyond the scope of this paper.

It is recognized that more nuanced investigations will be required before it is possible to answer completely the question of what mix matters. First, this study has investigated only one outcome, labour income, and there is widespread belief that different indicators of neighbourhood conditions may possess different explanatory power depending on which outcome is being considered (Brooks-Gunn *et al.*, 1997; Leventhal & Brooks-Gunn, 2000). Second, the study has investigated outcomes for adults only, and different household mix characteristics may prove more efficacious in predicting outcomes for children and youth (Ellen & Turner, 2003). Third, the analysis has considered additive, linear relationships between alternative indicators of household mix; more refined analyses should include checking for non-linearities in the neighbourhood effects and for interactions between an individual's characteristics (like educational or ethnic background) and the neighbourhood variables (Galster, 2003). It is hoped that the present paper's results are sufficiently provocative to stimulate such advances.

Notes

[1] For example, see the two volume *Neighbourhood Poverty*, (Eds) Brooks-Gunn *et al.* (1997); the reviews in Dietz (2002); Duncan & Raudenbush (1999); Ellen & Turner (1997, 2003); Friedrichs (1998); Friedrichs *et al.* (2003); Leventhal & Brooks-Gunn (2000); Robert (1999); Sampson *et al.* (2002); and the special issue of *Housing Studies*, 18(6), 2003.

[2] In Western Europe the issue links up with ongoing discussions over 'social exclusion'. Although not typically seen as a cause of social exclusion, it can serve as a mediator that can deepen or weaken it (Murie & Musterd, 2004). In the US the issue is linked with 'concentrated poverty', articulated most prominently by Wilson (1987) and Jargowsky (1997).

[3] For example, see Van Kempen (1997) and Bauder (2001). For a recent Swedish contribution to the study of spatial mismatch, see Åslund *et al.* (2007).

[4] The analysis intentionally excludes recent (after 1991) immigrants to Sweden because it is thought that their labour market experience is neither indicative of their longer-term economic value nor reflective of their initial neighbourhood environments when they enter Sweden. The authors are conducting a companion analysis that focuses on neighbourhood effects for immigrants.

[5] The log-linear transformation is not only appropriate given the positive skew of the income distribution, but also has sound grounding in economic theory, implicitly suggesting that income is an interactive (not additive) function of personal, neighbourhood, municipality and labour market characteristics.

[6] 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).

[7] 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 10 times as many people.

[8] Public rental in Sweden means almost entirely multi-family dwellings owned by municipal housing companies. These companies have emerged over a period of 60 years and they now possess about 20 per cent of all dwelling units (local variations). Public rental is not means tested and allocation is normally arranged in the form of waiting lists. Sweden does not have condominiums so owner occupation means single housing. Close to 50 per cent of the population live in owner-occupied houses. Urban residential segregation processes primarily sort people between home ownership and rental housing. See for example Musterd & Andersson (2005, note (i)) for more information on housing and residential segregation in Sweden.

[9] The control variables' coefficients vary little across the alternative neighbourhood mix specifications.

[10] R-squared value from a model with the control variables but without any neighbourhood variables is 0.455 for metropolitan males, 0.452 for metropolitan females, 0.459 for non-metropolitan males and 0.458 for non-metropolitan females.

References

Andersson, R. (2003) Urban development programmes in a Scandinavian welfare state: a top-down approach to bottom-up planning?, in: P. de Decker, J. Vranken, J. Beaumont & I. Nieuwenhuyze (Eds) *On the Origins of Urban Development Programmes in Nine European Countries*, pp. 165–181 (Antwerpen-Apeldoorn: Garant).

Andersson, R. (2006) Breaking segregation—Rhetorical construct or effective policy? The case of the metropolitan development initiative in Sweden, *Urban Studies*, 43, pp. 787–799.

Andersson, R. & Musterd, S. (2005) Area-based policies: a critical appraisal, *Tijdschrift voor economische en sociale geografie*, 96, pp. 377–389.

Åslund, O., Östh, J. & Zenou, Y. (2007) How important is access to jobs? Old question, improved answer, in: J. Östh (Ed.) *Home, Job and Space. Mapping and Modeling the Labor Market*. Geografiska Regionstudier No. 72 (Uppsala: Uppsala University).

Atkinson, R., Kintrea, K., Austin, M. D. & Baba, Y. (2001) Disentangling area effects: the contributions of place to household poverty, *Urban Studies*, 38, pp. 2277–2298.

Bauder, H. (2001) You're good with your hands, why don't you become an auto mechanic: neighborhood context, institutions and career development, *International Journal of Urban and Regional Research*, 25, pp. 593–608.

Beaumont, J., Burgers, J., Decker, K., Dukes, T., Musterd, S., Staring, R. & Van Kempen, R. (2003) Urban policy in the Netherlands, in: P. de Decker, J. Vranken, J. Beaumont & I. Nieuwenhuyze (Eds) *On the Origins of Urban Development Programmes in Nine European Countries*, pp. 119–137 (Antwerpen-Apeldoorn: Garant).

Belsley, D., Kuh, E. & Welsch, R. (1980) *Regression Diagnostics* (New York: Wiley).

Bergsten, Z. & Holmqvist, E. (2007) Blanda olika befolkningsgrupper fortfarande en målsättning—En undersökning av kommunala planeringsavdelningarnas och allmännyttiga bostadsföretags planeringsmål och strategier (Gävle: IBFs rapportserie).

Brewster, K. (1994) Race differences in sexual activity among adolescent women: the role of neighborhood characteristics, *American Sociological Review*, 59, pp. 408–424.

Brooks-Gunn, J., Duncan, G. J., Klebanov, P. K. & Sealand, N. (1993) Do neighbourhoods influence child and adolescent development?, *American Journal of Sociology*, 99, pp. 353–395.

Brooks-Gunn, J., Duncan, G. J. & Aber, J. L. (Eds) (1997) *Neighborhood Poverty: Vol. 1 Context and Consequences for Children* (New York: Russell Sage Foundation).

Cook, T. D., Shagle, S. C., Degirmencioglu, S. M., Coulton, C. J., Korbin, J. E. & Su, M. (1997) Capturing social process for testing mediational models of neighborhood effects, in: J. Brooks-Gunn, G. J. Duncan & L. Aber (Eds) *Neighborhood Poverty: Vol. 2 Policy Implications in Studying Neighborhoods*, pp. 94–119 (New York: Russell Sage Foundation).

Corcoran, M., Gordon, R., Laren, D. & Solon, G. (1987) Intergenerational transmission of education, income and earnings. Unpublished manuscript (Ann Arbor: University of Michigan).

Crane, J. (1991) The epidemic theory of ghettos and neighborhood effects on dropping out and teenage childbearing, *American Journal of Sociology*, 96, pp. 1226–1259.

Dawkins, C. J., Shen, Q. & Sanchez, T. W. (2005) Race, space, and unemployment duration, *Journal of Urban Economics*, 58, pp. 91–113.

Dietz, R. (2002) The estimation of neighborhood effects in the social sciences, *Social Science Research*, 31, pp. 539–575.

Duncan, G. J. (1994) Families and neighbors as sources of disadvantage in the school decisions of white and black adolescents, *American Journal of Education*, 103, pp. 20–53.

Duncan, G. J., Connell, J. P. & Klebanov, P. K. (1997) Conceptual and methodological issues in estimating causal effects of neighborhoods and family conditions on individual development, in: J. Brooks-Gunn, G. J. Duncan & J. L. Aber (Eds) *Neighborhood Poverty: Vol. 1 Context and Consequences for Children*, pp. 219–250 (New York: Russell Sage Foundation).

Duncan, G. & Raudenbush, S. (1999) Assessing the effect of context in studies of child and youth development, *Educational Psychology*, 34, pp. 29–41.

Ellen, I. & Turner, M. (1997) Does neighborhood matter? Assessing recent evidence, *Housing Policy Debate*, 8, pp. 833–866.

Ellen, I. & Turner, M. (2003) Do neighborhoods matter and why?, in: J. Goering & J. Feins (Eds) *Choosing a Better Life? Evaluating the Moving to Opportunity Experiment*, pp. 313–338 (Washington DC: Urban Institute Press).

Ensminger, M. E., Lamkin, R. P. & Jacobson, N. (1996) School leaving: a longitudinal perspective including neighborhood effects, *Child Development*, 67, pp. 2400–2416.

Friedrichs, J. (1998) Do poor neighbourhoods make their residents poorer? Context effects of poverty neighbourhoods on their residents, in: H. Andress (Ed.) *Empirical Poverty Research in a Comparative Perspective*, pp. 77–99 (Aldershot: Ashgate).

Friedrichs, J., Galster, G. & Musterd, S. (2003) Neighbourhood effects on social opportunities: the European and American research and policy context, *Housing Studies*, 18, pp. 797–806.

Galster, G. (2003) Investigating behavioural impacts of poor neighbourhoods: towards new data and analytic strategies, *Housing Studies*, 18, pp. 893–914.

Galster, G. (2005) *Neighbourhood Mix, Social Opportunities, and the Policy Challenges of an Increasingly Diverse Amsterdam* (Amsterdam: University of Amsterdam, Department of Geography, Planning, and International Development Studies). Available at: http://www.fmg.uva.nl/amidst/object.cfm/objectid = 7C149E7C-EC9F-4C2E-91DB7485C0839425.

Galster, G. (2007a) Neighbourhood social mix as a goal of housing policy: a theoretical analysis, *European Journal of Housing Policy*, 7, pp. 19–43.

Galster, G. (2007b) Should policymakers strive for neighbourhood social mix? An analysis of the Western European evidence base, *Housing Studies*, (forthcoming).

Galster, G. & Zobel, A. (1998) Will dispersed housing programmes reduce social problems in the US?, *Housing Studies*, 13, pp. 605–622.

Gephart, M. A. (1997) Neighborhoods and communities as contexts for development, in: J. Brooks-Gunn, G. J. Duncan & J. L. Aber (Eds) *Neighborhood Poverty: Vol. I Context and Consequences for Children*, pp. 1–43 (New York: Russell Sage Foundation).

Ginther, D., Haveman, R. & Wolfe, B. (2000) Neighbourhood attributes as determinants of children's outcomes: how robust are the relationships?, *Journal of Human Resources*, 35, pp. 603–642.

Goetz, E. G. (2002) *Clearing the Way* (Washington DC: Urban Institute Press).

Granovetter, M. (2005) The impact of social structure on economic outcomes, *Journal of Economic Perspectives*, 19, pp. 33–50.

Haurin, D. R., Dietz, R. D. & Weinberg, B. A. (2002) *The impact of neighborhood homeownership rates: a review of the theoretical and empirical literature*. Working paper (Columbus, OH: Department of Economics, Ohio State University).

Haveman, R. & Wolfe, B. (1994) *Succeeding Generations: On the Effects of Investments in Children* (New York: Russell Sage Foundation).

Hoynes, H. W. (2000) Local labor markets and welfare spells, *Review of Economics and Statistics*, 82, pp. 351–368.

Ionnides, Y. M. & Datcher Loury, L. (2004) Job information networks, neighborhood effects, and inequality, *Journal of Economic Literature*, 42, pp. 1056–1093.

Jargowsky, P. (1997) *Poverty and Place* (New York: Russell Sage Foundation).

Jenks, C. & Mayer, S. (1990) The social consequences of growing up in a poor neighborhood, in: L. E. Lynn & M. F. H. McGeary (Eds) *Inner-city Poverty in the United States*, pp. 111–186 (Washington DC: National Academy Press).

Johnson, M., Ladd, H. & Ludwig, J. (2002) The benefits and costs of residential mobility programmes for the poor, *Housing Studies*, 17, pp. 125–138.

Kearns, A. (2002) Response: from residential disadvantage to opportunity? Reflections on British and European policy and research, *Housing Studies*, 17, pp. 145–150.

Leventhal, T. & Brooks-Gunn, J. (2000) The neighborhoods they live in, *Psychological Bulletin*, 126, pp. 309–337.

Manski, C. F. (1995) *Identification Problems in the Social Sciences* (Cambridge, MA: Harvard University Press).

Manski, C. F. (2000) Economic analysis of social interactions, *Journal of Economic Perspectives*, 14, pp. 115–136.

Murie, A., Knorr-Siedow, T. & Van Kempen, R. (2003) *Large Housing Estates in Europe: General Developments and Theoretical Backgrounds* (Utrecht: Restate).

Murie, A. & Musterd, S. (2004) Social exclusion and opportunity structures in European cities and neighbourhoods, *Urban Studies*, 41, pp. 1441–1459.

Musterd, S. (2002) Response: mixed housing policy: a European (Dutch) perspective, *Housing Studies*, 17, pp. 139–144.

Musterd, S. & Andersson, R. (2005) Housing mix, social mix and social opportunities, *Urban Affairs Review*, 40, pp. 761–790.

Musterd, S. & Andersson, R. (2006) Employment, social mobility and neighbourhood effects, *International Journal for Urban and Regional Research*, 30, pp. 120–140.

Musterd, S., Ostendorf, W. & De Vos, S. (2003) Environmental effects and social mobility, *Housing Studies*, 18, pp. 877–892.

Ostendorf, W., Musterd, S. & de Vos, S. (2001) Social mix and the neighbourhood effect: policy ambition and empirical support, *Housing Studies*, 16, pp. 371–380.

Reardon, S. F. & Firebaugh, G. (2002) Measures of multigroup segregation, *Sociological Methodology*, 32, pp. 33–67.

Robert, S. (1999) Socioeconomic position and health: the independent contribution of community socioeconomic context, *Annual Reviews of Sociology*, 25, pp. 489–516.

Sampson, R. J. (1997) Collective regulation of adolescent misbehavior: validation results for eighty Chicago neighborhoods, *Journal of Adolescent Research*, 12, pp. 227–244.

Sampson, R. J., Morenoff, J. D. & Earls, F. (1999) Beyond social capital: spatial dynamics of collective efficacy for children, *American Sociological Review*, 64, pp. 633–660.

Sampson, R., Morenoff, J. & Gannon-Rowley, T. (2002) Assessing 'neighborhood effects': social processes and new directions in research, *Annual Reviews of Sociology*, 28, pp. 443–478.

South, S. & Crowder, K. (1999) Neighborhood effects on family formation: concentrated poverty and beyond, *American Sociological Review*, 64, pp. 113–132.

Theil, H. (1972) *Statistical Decomposition Analysis: With Applications in the Social and Administrative Sciences* (Amsterdam: North-Holland).

Van der Klaauw, B. & Van Ours, J. C. (2003) From welfare to work: does neighborhood matter?, *Journal of Public Economics*, 87, pp. 957–985.

Van Kempen, E. (1997) Poverty pockets and life chances, *American Behavioral Scientist*, 41, pp. 430–449.

Weinberg, B. A., Reagan, P. B. & Yankow, J. J. (2004) Do neighborhoods affect work behavior? Evidence from the NLSY79, forthcoming, *Journal of Labor Economics*, 24, pp. 891–924.

Wilson, W. J. (1987) *The Truly Disadvantaged* (Chicago, IL: University of Chicago Press).

Appendix

Table 1A. Regression results for control variables

Parameter	B	Std. Error	t
Intercept	6.6893	0.0569	117.52
No. children under age 7, 1995	−0.1873	0.0040	−46.66
Some sick leave during 1995 (1 = yes)	0.0880	0.0042	21.18
Pre-retired during 1995 (1 = yes)	−2.3630	0.0123	−192.44
Parental leave during 1995 (1 = yes)	0.1378	0.0052	26.63
Studying during 1995 (1 = yes)	−0.2201	0.0075	−29.32
No. years with pre-retirement, 1996–99 (1 = yes)	−0.5234	0.0030	−175.72
No. child-years under 7, 1996–99	−0.0862	0.0013	−65.70
No. years studying, 1996–99	−0.3575	0.0022	−159.96
No. years with parental leave, 1996–99	0.1439	0.0018	80.82
No. years with sick leave, 1996–99	0.1622	0.0015	110.26
Immigrants w/ < 5 years in Sweden (1 = yes)	−0.6511	0.0253	−25.75
Western country of birth (1 = yes)	−0.2763	0.0068	−40.73
Eastern European country of birth (1 = yes)	−0.5297	0.0123	−42.89
Non-Western country of birth (1 = yes)	−1.1387	0.0122	−93.56
No formal education (1 = yes)	−2.0446	0.0264	−77.48
< 10 years education (1 = yes)	−0.6330	0.0051	−123.61
10 years education (1 = yes)	−0.4369	0.0047	−92.49
13 years, some post-secondary (1 = yes)	0.3777	0.0091	41.59
14 + years, but no PhD (1 = yes)	0.4780	0.0039	122.76
PhD attained (1 = yes)	1.0229	0.0435	23.52
Single 1991 but couple 1995 (1 = yes)	−0.0434	0.0066	−6.59
Couple 1991 but single 1995 (1 = yes)	−0.1053	0.0061	−17.39
Education rose LT 11–12 to 11–12+(1 = yes)	0.4107	0.0262	15.66
Education rose 11–12 to higher (1 = yes)	0.4143	0.0123	33.70
Age in years	0.0016	0.0003	5.91
Age more than 50 years (1 = yes)	−0.2989	0.0052	−57.04
Single 1995 but couple 1999 (1 = yes)	−0.0922	0.0075	−12.24
Couple 1995 but single 1999 (1 = yes)	−0.1029	0.0063	−16.40
Mean income in local labour market, 1999	0.0003	0.0000	8.02

Note: [a]Results are shown for female-non-metropolitan stratum. The model also included the percentage of low-income males and the municipality dummies. All *p*-values are below 0.001.