

IZA DP No. 5238

Neighbourhood Choice and Neighbourhood Reproduction

Lina Bergström
Maarten van Ham
David Manley

October 2010

Neighbourhood Choice and Neighbourhood Reproduction

Lina Bergström

Uppsala University

Maarten van Ham

*University of St Andrews
and IZA*

David Manley

University of St Andrews

Discussion Paper No. 5238
October 2010

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Neighbourhood Choice and Neighbourhood Reproduction

Although we know a lot about why households choose certain dwellings, we know relatively little about the mechanisms behind neighbourhood choice. Most studies of neighbourhood choice only focus on one or two dimensions of neighbourhoods: typically poverty and ethnicity. This paper argues that neighbourhoods have multiple dimensions and that models of neighbourhood choice should take these dimensions into account. We propose the use of a conditional logit model. From this approach we can gain insight into the interaction between individual and neighbourhood characteristics which lead to the choice of a particular neighbourhood over alternative destinations. We use Swedish register data to model neighbourhood choice for all households which moved to a neighbourhood in the city of Uppsala between 1997 and 2006. Our results show that neighbourhood sorting is a highly structured process where households are very likely to choose neighbourhoods where the neighbourhood population matches their own characteristics.

JEL Classification: I30, J60, R23

Keywords: neighbourhood, housing choice, sorting, residential mobility, conditional logit, Sweden

Corresponding author:

Lina Bergström
Institute for Housing and Urban Research
Uppsala University
Box 785
SE-801 29 Gävle
Sweden
E-mail: lina.bergstrom@ibf.uu.se

Introduction

There is now increasing empirical evidence that characteristics of the residential context play a role in explaining moving behaviour (Clark et al, 2006; Feijten and van Ham, 2009; van Ham and Clark, 2009; van Ham and Feijten, 2008). Both the socioeconomic status of neighbourhoods (Harris, 1999) and the ethnic mix of the neighbourhood population (Clark, 1992; Crowder, 2000) have been linked to residential mobility decisions. Most studies investigating the effect of neighbourhood characteristics on mobility behaviour focus on the actual moving decision, or on the desire to move away from certain neighbourhoods. Much less is known about how households choose their destination neighbourhood, and which neighbourhood characteristics play a role for which types of households in these decisions.

Developing a better understanding of neighbourhood choice is beneficial for a number of linked literatures. The housing choice literature focuses mainly on characteristics of dwellings. These dwelling characteristics are not independent from neighbourhood choice as different neighbourhoods offer different types of dwellings. An improved understanding of neighbourhood choice is also crucial for our understanding of neighbourhood change. Changes in the demographic or socio-economic structure of neighbourhood populations can in part be explained through selective in- (and out-) flow of households. The literature on (the causes of) segregation will especially benefit from a better understanding of how households end up in certain neighbourhoods. Also the literature on neighbourhood effects will benefit from greater insight into the mechanisms which lead to selective inflows of households into neighbourhoods. The main challenge in the neighbourhood effects literature is to identify causal neighbourhood effects: selective inflows of households into neighbourhoods can bias models trying to measure neighbourhood effects (see, e.g., Bergström and van Ham, 2011; Duncan et al, 1997; Galster, 2008; van Ham and Manley, 2010).

Despite the centrality of neighbourhood choice in a number of literatures, our knowledge of the mechanisms behind neighbourhood choice is limited. Most studies of neighbourhood choice focus on whether or not a household ends up in a particular type of neighbourhood. These studies have two important limitations. The first is that the choice set of alternative neighbourhoods is not taken into account. The choice for a particular (type of) neighbourhood cannot be understood without knowledge about the potential alternatives. The second problem is that neighbourhoods are generally characterised based on a limited number of characteristics. Studies typically model whether households move into a poverty neighbourhood or not, or into an ethnic concentration neighbourhood or not. Some studies characterise neighbourhoods based on two characteristics (typically combinations of poverty and ethnicity). The choice of a particular neighbourhood is often much more complicated than this basic conceptualisation allows (Bergström and van Ham, 2011). These two limitations are strongly related to the types of models used in studies of neighbourhood choice: typically binary or multinomial logit models which only allow for a limited number of outcome categories (neighbourhood types).

This paper contributes to the literature on neighbourhood choice by explicitly taking choice sets and multiple neighbourhood characteristics into account when modelling neighbourhood choice. We propose the use of a conditional logit model which allows us to investigate the choice of a specific neighbourhood (not a neighbourhood type), while taking multiple neighbourhood characteristics into account. The conditional logit model has rarely been used in studies of neighbourhood

choice (exceptions are Ioannides and Zabel, 2008; Quillian and Bruch, 2010). We use data from the GeoSweden database, a longitudinal register-based database with information on all individuals who has resided in Sweden since 1990. For the purpose of our study we have selected all households which moved to or within the city of Uppsala, a medium-sized city in the mid-east of Sweden, during a ten-year period (1997-2006).

Literature review

The process leading to residential mobility is often conceptualised as two separate but related decisions; the choice of whether to move or not (sometimes subdivided in moving desires, intentions and plans), and (following that initial decision), the choice of destination (Brown and Moore, 1970). The primary driver of these decisions is household dissatisfaction with the existing dwelling in relation to household preferences and needs (Brown and Moore, 1970; Clark and Dieleman, 1996). The role of neighbourhood characteristics in mobility decisions has received less attention, but the body of literature which includes neighbourhood characteristics in models of residential mobility is growing (Clark and Ledwith, 2006; Clark et al, 2006; Feijten and van Ham, 2009; Kearns and Parkes, 2003; Lee et al, 1994; Permentier et al, 2009; van Ham and Clark, 2009; van Ham and Feijten, 2008). Studies investigating the second step in the decision making process (the choice of destination) have mainly focused on dwelling types (tenure, size, price) and less attention has been paid to location and neighbourhood characteristics. The choice of a particular dwelling in a particular neighbourhood is based on household preferences, resources and restrictions, and on the availability of dwellings (van Ham, forthcoming). It can be argued that the choice of a particular neighbourhood is (largely) the result of the desire to live in a specific type (tenure, size, and price) of dwelling. There is large variation in the types (and quality) of dwellings available between neighbourhoods. However, there are clear indications in the literature that households also take neighbourhood attributes such as reputation, status, population composition, and location into consideration when choosing a dwelling (see for example Michelson, 1977; Permentier et al., 2009).

Households can only choose to live in a particular neighbourhood when there are vacancies available. Bailey and Livingstone (2007) have shown that the primary drivers of vacancy rates are demographic mix and housing stock characteristics (see also van Ham and Clark, 2009). Neighbourhoods dominated by younger age groups, single adults, and with a high share of dwellings in the rental sectors are likely to have the highest turnover rates. Such neighbourhoods are also associated with lower income groups. Work by Andersson and BråmÅ (2004) has shown that neighbourhood mobility in Sweden is greatest in neighbourhoods with high proportions of non-native residents (see van Ham and Clark, 2009 for similar evidence from the Netherlands).

Financial resources are an important driver of neighbourhood choice. Buying a house, or buying into a housing cooperative in Sweden, requires substantial down payments, making neighbourhoods where such dwellings dominate virtually inaccessible to households with limited resources. The Swedish housing market is strongly segmented, leading to an uneven distribution of socio-economic and ethnic groups within different types of housing and tenure. For instance, immigrants from non-western countries are over-represented in the rental market, not necessarily as the result of free choice, but as a result of a lack of resources (Andersson et al, 2007;

Murdie and Borgegård, 1998). Ethnic minorities often live concentrated in a limited number of neighbourhoods because different housing types and tenures are unevenly distributed within the urban space. Large groups are excluded from certain parts of cities simply because they cannot afford to live there. Previous Swedish studies have emphasized the role of (spatial) segmentation when explaining ethnic residential segregation. As expressed by Andersson and colleagues (2007, p.5), “the geographic distribution of tenure forms and housing types within a city or region is [...] a fundamental condition for the segregation process”. For the Swedish rental segment, where prices are set according to size and quality rather than direct market value, this analysis is less valid. Nevertheless, there are alternative costs associated with the more desirable dwellings and locations in the rental market. For instance, the attractiveness of the dwelling and neighbourhood is reflected in the waiting time required to successfully secure a dwelling. The waiting time for a dwelling in the most popular areas of larger cities is over 10 years. Households with urgent housing needs are often forced to accept a dwelling in less attractive neighbourhoods.

The international literature emphasises the effect of income on neighbourhood sorting. For instance, Clark and Ledwith (2006) analysed moving desires and actual moves in Los Angeles, using longitudinal data collected in 2002. They concluded that the greatest driver of neighbourhood sorting for both Whites and Hispanics remains household income. Other studies have emphasized how limited financial resources prevent people from exiting areas with high levels of poverty (Bolt and van Kempen, 2003; Kearns and Parkes, 2003; South and Crowder, 1997; 1998). We simply note that households with fewer available resources will have more restricted choice sets, and that some households will have very limited choice sets. We thus *hypothesize that households sort into neighbourhoods according to socio-economic characteristics (income, education, employment status, use of benefits) (Hypothesis 1)*.

Neighbourhood choice has been associated with the ethnic composition of the neighbourhood population. Starting with the work of Schelling (1969; 1971) it has been argued that many households have a preference for living in neighbourhoods with households of similar ethnic background and that these preferences can lead to highly segregated neighbourhoods. Clark (1991) used data for the United States to demonstrate that the Schelling hypothesis was broadly correct, and that even small differences in preferences between ethnic groups with regard to the ethnic composition of neighbourhoods can lead to highly segregated communities (see also Fossett, 2006). Evidence from both Europe and the United States indicate that ethnic segregation is primarily driven by own-group preferences held by the majority population. In addition, the majority population tend to have the greatest level of resources and therefore, the ability to put these preferences in to action. For example, Clark (1991) reports that while whites preferred the ethnic mix in their neighbourhood to be at least 80% white, blacks seemed to prefer a 50/50 mix. In Sweden, most immigrant-dense neighbourhoods are truly multicultural, making the notion of voluntary ethnic minority clustering unlikely. Ethnic mix preferences (or rather preferences for relative homogeneity) by the majority population is apparent at the aggregate level through patterns of white avoidance of ethnic minority neighbourhoods (BråmÅ, 2006).

Studies from various countries have found that ethnic minorities are more likely than natives to move to ethnic concentration neighbourhoods. It has been hypothesised that these moves to ethnic concentration neighbourhoods are (partly) motivated by the desire to live in areas with others who have common life experiences and by the availability of ethnic specific services (see Bowes et al, 1997). Other

studies have emphasized the impact of socio-economic differences between ethnic and non-ethnic groups (e.g. Clark and Ledwith, 2006; South and Crowder, 1997; 1998). In Sweden ethnic minorities are overrepresented among the lower-income groups, and as a result they concentrate in less low cost neighbourhoods (see Bolt et al, 2008, for a similar explanation for the Netherlands). Bråmås and Andersson (2005; 2010) have shown that recent immigrants in Sweden initially move to areas with high densities of immigrants. When their income improves they are more likely to leave these neighbourhoods and move to less ethnically concentrated neighbourhoods. Discrimination has also been shown to influence neighbourhood ethnic sorting (e.g. Turner et al, 2002, for the United States), although the extent to which this is valid for Sweden is unclear (Bråmås, 2007). Based on the above literature review, we *hypothesise that households sort into neighbourhoods based on their ethnicity, with foreign born households being more likely to move to immigrant-dense neighbourhoods compared to native Swedes (Hypothesis 2).*

There is a large literature showing how the propensity to move and the choice of dwellings is affected by life course factors. The literature emphasises the relationships between housing choice and the need for space and household and socio-economic change over the life course (see Clark and Dieleman, 1996; Clark and Huang, 2003). Changes in household composition are often associated with moves to different types of dwellings in different types of neighbourhoods. Preferences, resources and constraints all change over the life course. In general, young adults and singles prefer to live in city centres, while households with children prefer to live in suburbs. Households with children generally prefer to live in a neighbourhood with many other children, with good access to (high quality) services such as schools and playgrounds, and with dwellings that match the needs of the family (see Clark and Dieleman, 1996). Based on the above we *hypothesise that households sort into neighbourhoods according to demographic characteristics, for example, households with children are more likely to move to neighbourhoods with a high proportion of households with children (Hypothesis 3).*

The literature review has highlighted a range of factors which influence neighbourhood choice. Most of these factors have been studied in isolation while in reality households ‘buy’ a bundle of neighbourhood attributes when moving to a particular neighbourhood. The literature review suggests that neighbourhood sorting is a highly structured process where households generally move to neighbourhoods with populations similar to their own characteristics.

Method and data

The majority of studies investigating neighbourhood choice employ binary logistic regression or generalized multinomial logit models. The disadvantage of these types of models is that they only allow for a restricted number of neighbourhood characteristics to be taken into account: typically the level of deprivation and/or the level of concentration of ethnic minorities in the neighbourhood. This highly reductionist approach limits our understanding of how households choose neighbourhoods as neighbourhoods have multiple dimensions, and these dimensions can be expected to interact with household characteristics.

In this paper we adopt a statistical model that allows for both large choice sets and multiple characteristics of destinations. We model neighbourhood choice using a conditional logit model (CLM) (McFadden, 1974; see also Hoffman and Duncan,

1988, which includes a comparison to multinomial logit models). The CLM has been widely used in biomedical research, but also in the social sciences, especially in studies of transportation demand and consumer choices. Few previous studies have used CLM to investigate residential mobility and neighbourhood choice (notable exceptions are Ioannides and Zabel, 2008; Quillian and Bruch, 2010). The CLM is closely related to the generalized multinomial model but where the multinomial model treats neighbourhood choice as a function of household¹ characteristics, the CLM treats neighbourhood choice as a function of the characteristics of the alternatives (potential destinations). The comparison in a CLM is made *within* rather than *between* households, as would be the case in a multinomial logit, estimating the probability that household i will choose alternative j among a set of alternative neighbourhoods. The choice set of neighbourhoods in a CLM can theoretically be of any size, and include all neighbourhoods within a city or a housing market. The impact of the characteristics of household i is assumed to be constant across alternative neighbourhoods. Thus, let P_{ij} denote the probability that household i will choose alternative j . N_j refers to the characteristics of the j^{th} alternative, and $C_{(i)}$ the set of alternatives (the choice set) for the i^{th} household. Following Hoffman and Duncan (1998), the conditional logit model is written:

$$P_{ij}(N_j C_{(i)}) = \exp(\beta N_{ij}) / \sum_{k=1}^K \exp(\beta N_{ik}) \quad (1)$$

So the CLM estimates the probability that a household chooses a particular neighbourhood from a choice set of neighbourhoods, based on neighbourhood characteristics. When modelling neighbourhood choice it is also important to take household characteristics into account. Since household characteristics are constant within the choice set of a household, household characteristics must be interacted with neighbourhood characteristics to be included in the model. Let X_i denote the characteristics of the i^{th} household.

$$P_{ij}(N_j, X_i, C_{(i)}) = \exp(\beta N_j X_i) / \sum_{k=1}^K \exp(\beta N_k X_i) \quad (2)$$

Both household and neighbourhood characteristics are measured at time $t-1$, the year before the actual move took place. This is important to identify the causal path between household and neighbourhood characteristics and actual moves. If characteristics were measured at time t , it would be impossible to determine whether households' characteristics were the cause or result of a move. The exception to this are newly formed households (after separation, couple formation, or children leaving the parental home, 21 % of the moving households), where characteristics are measured at time t (the formation of the household was most likely the reason for the move). The probability that the i^{th} household will choose the j^{th} alternative, or in other words, will live in neighbourhood j at time t is thus written:

$$P_{ijt}(N_{j(t-1)}, X_{i(t-1)}, C_{(i)}) = \exp(\beta N_{j(t-1)} X_{i(t-1)}) / \sum_{k=1}^K \exp(\beta N_{k(t-1)} X_{i(t-1)}) \quad (3)$$

¹ In terms of neighbourhood choices, it can be argued that decisions are made at a household level, rather than at the level of the individual.

The data used in this study are obtained from the GeoSweden database, a longitudinal register-based database with information on all individuals who have resided in Sweden at some point since 1990. The data is unique in the sense that it includes the total Swedish population. It is collected by Statistics Sweden from various public records and is updated annually. The GeoSweden database includes demographic, socio-economic and geographic information (geo-coded to a 100 by 100 meter grid) which makes it possible to identify and track movers. For this study we have aggregated the individual level data in GeoSweden to a household level. We selected all households that had moved either within or into the city of Uppsala between 1997 and 2006. Those who left Uppsala are excluded from the dataset. A move is defined as a change of 100 m² grid between two consecutive years. The ten years of data correspond to nine years of moving data, 1997/1998 through to 2005/2006, which provides a total of 130,280 household moves.

Although the GeoSweden data is unusually rich for register data, there are some limitations. The most important limitation is that couples can only be identified in the dataset if they are married, registered partners or have a dependent child in common. As a result, cohabiting couples without children will be observed as (two) single households. Since cohabiting is common in Sweden, especially in the young (and mobile) age groups, the number of moving households will be exaggerated (the data contains too many moving singles). Another limitation of the data is that it only allows the identification of employed individuals. The group of non-employed individuals is very heterogeneous and includes the unemployed, students, retired people and other non-employed individuals. We have defined unemployment as individuals in working ages (20-64) that are not employed or do not receive student funding. Notwithstanding these limitations, we feel that the quality, breadth and extent of the data is sufficient to enable meaningful analysis.

The most appropriate definition of neighbourhoods in studies of neighbourhood choice is highly contested (see discussions by Galster, 2001; Kearns and Parkinson, 2001; Manley et al, 2006). Several different spatial levels have been used to identify neighbourhoods in the literature but the most common are census tracts, postal codes, or equivalent administrative divisions. We make neighbourhoods operational using the Swedish SAMS (Small Area Market Statistics) areas. These have been used to approximate neighbourhoods in previous studies using the same GeoSweden database (Andersson and Bråmås, 2004; Bråmås and Andersson, 2005; 2010; Galster et al, 2008). We have excluded all SAMS (herein referred to as neighbourhoods) from the choice set of neighbourhoods with a population of less than 50 inhabitants (mostly commercial and industrial areas). We have also removed all moves into these neighbourhoods. The total number of neighbourhoods included in the analysis in 2001 (the midyear of the period under investigation) is 122. They contain 1103 individuals on average (see Table 1). Since neighbourhood populations change between years, in some cases dramatically as a result of new dwelling construction or old dwelling demolition, the number of neighbourhoods included in the choice sets varies between the years, from 120 (in 2000) to 123 (in 2004 and 2005). We have included moves between and within neighbourhoods, which resulted in a total of 130,231 moves in our dataset.

For each moving household we constructed a choice set of neighbourhoods in Uppsala. Theoretically, this could consist of all neighbourhoods in the city, which in 2001, for example, would mean 122 neighbourhoods. Because the dataset for a CLM contains a line of data for each neighbourhood in the choice set, including the neighbourhood entered, we would have ended up with over 15 million lines of data

(130,231 moves each having approximately 122 neighbourhoods to choose from). Given that we included 8 neighbourhood characteristics and a large range of household characteristics, the final dataset was too large to work with given software and hardware restrictions. Following Ioannides and Zabel (2008), we have therefore chosen to limit the choice set to ten neighbourhoods per household. This set includes the destination neighbourhood and nine randomly chosen alternative neighbourhoods in Uppsala. This resulted in a dataset with around 1.3 million lines². To account for the unequal distributions of moving opportunities between neighbourhoods – in larger neighbourhoods it is more likely that a vacancy becomes available during a given time span than in smaller neighbourhoods – the alternatives in the choice set are weighted according to the number of vacancies (defined by the number of households moving out of the neighbourhood) during the year of the move, using the total number of vacancies over all neighbourhoods in the particular year as denominator. The CLM we use therefore estimates the probability that a household chooses a neighbourhood out of a set of 10 alternatives.

The CLM includes a set of household and neighbourhood characteristics as independent variables. All household level variables (ethnic composition, family composition, whether there is a student in the household, highest level of education, employment status, family disposable income, benefit status, and tenure) in the model are transformed into dichotomous categories. Ethnic composition can be Swedish, if all adult members are Swedish born, foreign born, if all adult members are foreign born, or mixed. Employment status is calculated in a similar way. Education level is the highest level in the household, student is coded as 1 if any of the adult household members are students, while the remaining variables directly relate to households in the GeoSweden data. Neighbourhood level variables are included as the proportion of the population belonging to that group. The only exception is median neighbourhood income, which is included as a continuous variable. We have then interacted each set of household variables with their equivalent variable on the neighbourhood level. For example, the median neighbourhood income is interacted with each of the five income quintiles on the household level, while the proportion of inhabitants that are foreign born in the neighbourhood is interacted with the ethnic categories on the household level (all adults born in Sweden, all adults being foreign born, or mixed).

Table 1 presents the variable summary statistics for neighbourhood characteristics in 2001 (the midyear of the period under investigation), including the proportion of foreign born, households with children under age 18, students living in the neighbourhood, people with low education (no high school degree), non-employed (including individuals in working ages that are neither employed nor students), people receiving social benefits, people living in public rented dwellings, and median household disposable income (continuous). Table 2 presents summary statistics for the characteristics of moving households and all households who lived in Uppsala in 2001. The Table shows that foreign born residents, singles, students, non-employed, receivers of benefits are overrepresented among movers.

<<<<<Table 1 about here>>>>>

² We have also run the model with a choice set including all 120 neighbourhoods in Uppsala that fulfil our population requirements during the entire period, using a random 25 % sample of the population of movers. The results are very similar to models using a randomized choice set of 10 neighbourhoods but including all moving households, but the explanatory power of the model is reduced. A probable explanation is that a city contains many neighbourhoods which are very similar in terms of their characteristics, which reduces the power of the model.

<<<<Table 2 about here>>>>

Results

To get an idea of the distribution of vacancies in Uppsala, we calculated the share of households moving within or away from their neighbourhood (in 2001). In Uppsala, turnover rates vary from 0 to 46.5 % with a mean of 15 %, meaning that about 1/7 of the population in an average neighbourhood moved during 2001. In Table 3 we show the characteristics of neighbourhood populations for 5 types of neighbourhoods, based on quintiles of population turnover rates. Neighbourhoods in the highest turnover quintile group (quintile 5) have a mean turnover rate of 26.8 %. Further investigation reveals that the most mobile neighbourhoods are typically student areas; the proportion of students in these neighbourhoods is almost twice as high as in the areas in the second highest turnover quintile (quintile 4).

<<<<Table 3 about here>>>>

Table 3 also shows almost linear correlations between turnover rates and a number of neighbourhood characteristics: the higher the share of foreign born, non-employed people, and people on social benefits, and the lower the share of households with children, and median income in the neighbourhood, the higher turnover rates are. The correlations between turnover rate, income and ethnicity are shown in Figure 1, where each variable above has been transformed into quintiles. The darker shaded areas are neighbourhoods where turnover rate (1a), median income (1b) or the share of foreign born (1c) is high.

<<<<Figure 1 about here>>>>

Table 4 presents the results of the conditional logit model of neighbourhood choice. In Model 1 we only included the interaction effect between median neighbourhood income and household income in quintiles (the 1st quintile includes households with the lowest incomes). In the literature review we hypothesized income to be a major driver of neighbourhood sorting because income will be the main constraint regarding access to better neighbourhoods. We find that with increasing income in the neighbourhood, households with the lowest incomes are the least likely to move there, and households with the highest incomes are the most likely to move there. The interaction effect of neighbourhood and household income explains over 12 percent of the total variation in neighbourhood choice, which is relatively high since Model 2 – including a range of other interaction effects – explains just over 18 percent.

<<<<Table 4 about here>>>>

Model 2 shows that after including other interaction effects, the gap between the income groups is somewhat reduced compared to Model 1. Nevertheless, the differences between the household income categories remain significant and show the same linear pattern. Also the other interaction effects demonstrate that households are likely to choose neighbourhoods with populations similar to themselves, and avoid neighbourhoods where inhabitants are different. The interaction effect between the share of foreign born people in the neighbourhood and household ethnic composition

shows that households with only Swedish born people are the least likely to move to a neighbourhood with a high share of foreign born. The foreign born (households where all adult members were born outside Sweden) are more likely to move to a neighbourhood with a high share of foreign born. As would be expected given these results, mixed ethnic households are between these two groups. These results confirm that neighbourhood sorting based on ethnic background takes place in Sweden. The underlying mechanism is very likely to be choice (or discrimination) as we control for other neighbourhood characteristics such as income. The results also show how ethnic segregation is caused and maintained through selective neighbourhood sorting by ethnicity.

Families with one or more children are the most likely of all household categories to move to neighbourhoods with a high percentage of households with children. Singles without children are the least likely to move to neighbourhoods with a high percentage of households with children. Student households (where at least one member of the household is registered as a student) are more likely to move to neighbourhoods with a high percentage of students than households without any students. Households without any students seem to avoid student areas. It should be noted that in some neighbourhoods high concentrations of student accommodation can be found. These neighbourhoods are less likely to attract other types of households because the accommodation is neither suitable nor available for them. The effect of the housing market will be further discussed in the conclusion of this paper.

Households without anyone with a high school degree are the most likely to move to neighbourhoods with a high share of people with a low level of education. Households with at least one person with a university degree are the least likely to move to neighbourhoods with a high percentage of people with a low level of education. A similar pattern can be observed for the interaction effect between the neighbourhood level of non-employment and the employment status of households (households being fully employed, fully unemployed, mixed, or where the members are of age 65 or older). Although all coefficients are negative and the differences are very small, we find that employed households are less likely to move to neighbourhoods with a high proportion of unemployed individuals. Likewise, households not in receipt of social benefits avoid neighbourhoods with a high proportion of households on benefits. The coefficients for the tenure interactions are all very small, indicating that the tenure a household resides in prior to the move has little influence on the choice of neighbourhood, after controlling for all the other interactions. The results show a small tendency for public renters to move to neighbourhoods which already have a high share of public rental residents. Conversely, home owners are those most likely to avoid public rental neighbourhoods, followed by households living in cooperatives and private renters.

Our results show clear indications of a structured neighbourhood sorting process. Households consistently choose neighbourhoods that match their own characteristics, and they do so for all neighbourhood characteristics included in Model 2. Although some variables appear to have a small impact on the sorting process, especially in comparison to income, the patterns are consistent. As Clark (1992) noted when reporting on an empirical investigation of Schelling's hypothesis (1969; 1971), small differences in preferences with regard to the neighbourhood population composition can lead to marked segregation outcomes in neighbourhoods over time. The results strongly suggest that neighbourhood characteristics are reproduced through selective mobility of households.

Discussion

This paper contributes to the literature on neighbourhood choice by explicitly taking choice sets and multiple neighbourhood characteristics into account when modelling neighbourhood choice. This is one of the first studies of neighbourhood choice to use a conditional logit model. The main advantage of this model is that it allowed us to investigate the choice for a specific neighbourhood, while taking multiple neighbourhood characteristics into account. The modelling results indicate that moving patterns are highly structured and that households are likely to move to neighbourhoods where the population has similar characteristics than themselves (supporting hypotheses 1, 2 and 3). We have found this to be valid for all combinations of household and neighbourhood characteristics included in the model. The use of a conditional logit model allowed us to control the effect of each neighbourhood characteristics for all other characteristics. Household income seems to be the key characteristic influencing neighbourhood sorting in Sweden.

The results clearly indicate that the neighbourhood choices made by moving households result in the reproduction of neighbourhood characteristics over time. It is important here to stress that the concept of choice needs to be used with caution. Households make choices within a restricted choice set. Choices are restricted by household preferences, resources and restrictions, but also by constraints imposed by the structure of the housing market. It is very likely that poor households do not 'choose' to move to poverty neighbourhoods, but move there because they cannot afford to live anywhere else (van Ham and Manley, 2009; 2010). Their level of choice is seriously reduced due to a lack of realistic alternatives. For other combinations of household and neighbourhood characteristics a range of mechanisms is most likely at play. Households with children move to neighbourhoods with a high share of households with children because that is where suitable dwellings can be found, but also because they prefer to live in a child friendly neighbourhood. The effect of ethnicity is even more complicated as preferences, discrimination and the structure of the housing market are all likely to play a role. Although the fact that we controlled for tenure and neighbourhood income suggest that preferences and discrimination are the most important drivers of the segregation process.

Our findings on neighbourhood sorting are beneficial for literatures on housing choice, neighbourhood change, segregation, and neighbourhood effects. The neighbourhood effects literature can benefit from a more in-depth understanding of sorting processes as selective mobility into neighbourhoods hinders the identification of causal neighbourhood effects (Bergström and van Ham, 2011). To assess whether neighbourhood deprivation has a causal effect on individual outcomes, such as employment outcomes, it is crucial to take into account that unemployed people are more likely to move to deprived neighbourhoods in the first place (see van Ham and Manley, 2010). In this paper we have shown that residential mobility is highly selective, that neighbourhood sorting is highly structured, and that mobility is highest in the deprived neighbourhoods that tend to be the focus of much neighbourhood effects research. Our results indicate that people living in poverty areas were likely to be poor at the time of entry into these neighbourhoods. This seriously undermines any claim of causality when studying the effects of living in deprived neighbourhoods on individual poverty. Furthermore, the high turnover rates in distressed areas mean that exposure times in these neighbourhoods tend to be short and that the neighbourhoods do not provide stable environments even for those staying put. High mobility rates

mean that neighbourhood transmission mechanisms that require a certain amount of exposure to the same people are less likely to take place (see for instance Bergström, forthcoming; Galster, 2011). The results of this study should lead to better models of neighbourhood effects, explicitly taking neighbourhood sorting into account.

Sweden is internationally known for being a relatively 'equal' society, whether it comes to income or opportunities. The Swedish housing market further enhance this picture, at least when it comes to the rental housing market, as the public rental segment is open to everyone, is relatively attractive compared to the social housing estates and poverty neighbourhoods found in parts of Europe and the U.S., and rent levels are not market driven. Nevertheless, we have found clear evidence of neighbourhood sorting according to demographic, socio-economic and ethnic dimensions in Sweden, and this suggests that neighbourhood sorting is structural within housing market and cannot be ignored.

Acknowledgements

This paper was written whilst Lina Bergström was a visiting researcher at the Centre for Housing Research (CHR) at the University of St Andrews. We acknowledge the financial support of CHR which made this stay possible.

References

- Andersson R, BråmÅ Å, 2004, "Selective Migration in Swedish Distressed Neighbourhoods: Can Area-based Urban Policies Counteract Segregation Processes?" *Housing Studies* **19** 517-539
- Andersson R, Hogdal J, Johansson S, 2007, *Planering för minskad boendesegregation*, Rapport 2007:1, Regionplane- och trafikkontoret (Stockholms läns landsting: Stockholm)
- Bailey N, Livingston M, 2007, *Population Turnover and Area Deprivation* (Joseph Rowntree Foundations: Bristol)
- Bergström L, (forthcoming), "The Impact of Residential Mobility on Measurements of Neighbourhood Effects" *Housing Studies*
- Bergström L, van Ham M, 2011, "Understanding neighbourhood effects: selection bias and residential mobility", in *Neighbourhood Effects Research: New Perspectives* Eds M van Ham, D Manley, N Bailey, L Simpson, D Maclennan (Dordrecht: Springer)
- Bolt G, van Kempen R, 2003, "Escaping Poverty Neighbourhoods in the Netherlands" *Housing, Theory and Society* **20** 209-222
- Bolt G, van Kempen R, van Ham M, 2008 "Minority Ethnic Groups in the Dutch Housing Market: Spatial Segregation, Relocation Dynamics and Housing Policy" *Urban Studies* **45** 1359-1384
- Bowes A, Dar N, Sim D, 1997, "Tenure preference and housing strategy: an exploration of Pakistani experiences" *Housing Studies* **12** 63-84
- BråmÅ Å, 2006, "'White flight'? The production and reproduction of immigrant concentration areas in Swedish cities, 1990 – 2000" *Urban Studies* **43** 1127-1146
- BråmÅ Å, 2007, *Etnisk diskriminering på bostadsmarknaden. En forskningsöversikt*, IBF Working paper no 54 (Gävle: Institute for Housing and Urban Research)

- Bråmås Å, Andersson R, 2005, "Who leaves Sweden's large housing estates?" in *Restructuring large housing estates in Europe* Eds R van Kempen, K Dekker, S Hall, I Tosics (The Policy Press: Bristol)
- Bråmås Å, Andersson R, 2010, "Who leaves rental housing? Examining possible explanations for ethnic housing segmentation in Uppsala, Sweden" *Journal of Housing and the Built Environment*, Published online: DOI: 10.1007/s10901-010-9179-4.
- Brown L A, Moore E G, 1970, "The Intra-Urban Migration Process: A Perspective" *Geografiska Annaler B* **52** 1-13
- Clark W A V, 1991, "Residential Preferences and Neighborhood Racial Segregation: A Test of the Schelling Segregation model" *Demography* **28** 1-19.
- Clark W A V, 1992, "Residential preferences and residential choices in a multiethnic context" *Demography* **29** 451-466
- Clark W A V, Deurloo M, Dieleman F, 2006, "Residential Mobility and Neighbourhood Outcomes" *Housing Studies* **21** 232-342
- Clark, W A V, Dieleman F M, 1996, *Households and Housing: Choice and Outcomes in the Housing Market* (CUPR Press: New Brunswick)
- Clark W A V, Huang Y, 2003, "The life course and residential mobility in British Housing markets" *Environment and Planning A* **35** 323-339
- Clark W A V, Ledwith V, 2006, "Mobility, housing stress and neighbourhood contexts: evidence from Los Angeles" *Environment and Planning A* **38** 1077-1093
- Crowder K, 2000, "The racial context of white mobility: an individual-level assessment of the white flight hypothesis" *Social Science Research* **29** 223-257
- Duncan G, Connell J, Klebanov P, 1997, "Conceptual and methodological issues in estimating causal effects of neighborhoods and family conditions on individual development" in *Neighborhood Poverty: Vol. 1 Context and Consequences for Children* Eds J Brooks-Gunn, G Duncan, J Aber (Russell Sage Foundation: New York)
- Feijten P, van Ham M, 2009, "Neighbourhood change ... reason to leave?" *Urban Studies* **46** 2103-2122
- Fosset M, 2006, "Including Preference and Social Distance Dynamics in Multi-Factor Theories of Segregation" *Journal of Mathematical Sociology* **30** 289-298
- Galster G, 2001, "On the Nature of Neighborhood" *Urban Studies* **38** 2111-2124
- Galster G C, 2008, "Quantifying the Effect of Neighbourhood on Individuals: Challenges, Alternative Approaches, and Promising Directions" *Journal of Applied Social Sciences, Schmollers Jahrbuch* **128** 1-42
- Galster G, 2011, "The Mechanism(s) of Neighborhood Effects. Theory, Evidence, and Policy Implications", in *Neighbourhood Effects Research: New Perspectives* Eds M van Ham, D Manley, N Bailey, L Simpson, D Maclennan (Dordrecht: Springer)
- Galster G, Andersson R, Musterd S, Kauppinen T, 2008, "Does Neighborhood Income Mix Affect Earnings of Adults? New Evidence from Sweden" *Journal of Urban Economics* **63** 858-870
- Harris D R, 1999, "'Property values drop when blacks move in, because ...': racial and socioeconomic determinants of neighborhood desirability" *American Sociological Review* **64** 461-479
- Hoffman S D, Duncan G J, 1988, "Multinomial and Conditional Logit Discrete-Choice Models in Demography" *Demography* **25** 415-427

- Ioannides Y M, Zabel J E, 2008, "Interactions, neighborhood selection and housing demand" *Journal of Urban Economics* **63** 229-252
- Kearns A, Parkes A, 2003, "Living in and leaving poor neighbourhoods in England" *Housing Studies* **18** 827-851
- Kearns A, Parkinson M, 2001, "The significance of neighbourhood" *Urban Studies* **38** 2103-2110
- Lee B A, Oropresa R S, Kanan J W, 1994, "Neighborhood Context and Residential Mobility" *Demography* **31** 249-270
- Manley D, Flowerdew R, Steel D, 2006, "Scales, levels and processes: studying spatial patterns of British census variables" *Computers, Environment and Urban Systems* **30** 143-160
- McFadden D, 1974, "Conditional logit analysis of qualitative choice behaviour", in *Frontiers in econometrics* Ed P Zarembka (Academic Press: New York)
- Michelson W, 1977, *Environmental choice, human behavior, and residential satisfaction* (Oxford University Press: New York)
- Murdie R A, Borgegård L-E, 1998, "Immigration, Spatial Segregation and Housing Segmentation of Immigrants in Metropolitan Stockholm, 1960-95" *Urban Studies* **35** 1869-1888
- Permentier M, van Ham M, Bolt G, 2009, "Neighbourhood reputation and the intention to leave the neighbourhood" *Environment and Planning A* **41** 2162-2180
- Quillian L, Bruch E, 2010, *Race and Class in Neighborhood Mobility: A Conditional Logit Model of Neighbourhood Migration* Paper presented at the meeting of the Population Association of America, Dallas, Texas, April 2010
- Schelling T C, 1969, "Models of segregation" *The American Economic Review* **59** 488-93
- Schelling T C, 1971, "Dynamic models of segregation" *Journal of mathematical Sociology* **1** 143-186
- South S J, Crowder K D, 1997, "Escaping Distressed Neighbourhoods: Individual, Community and Metropolitan Influences" *American Journal of Sociology* **102** 1040-4084
- South S J, Crowder K D, 1998, "Leaving the 'Hood': Residential Mobility between Black, White and Integrated Neighbourhoods" *American Sociological Review* **63** 17-26
- Turner M A, Ross S L, Galster G C, Yinger J, 2002, *Discrimination in Metropolitan Housing Markets: National Results from Phase I* HDS 2000 (The Urban Institute: Washington D.C)
- van Ham M, forthcoming, "Housing Behaviour" in *Handbook of Housing Studies* Eds D Clapham, W A V Clark, K Gibb (Sage: London)
- van Ham M, Clark W A V, 2009, "Neighbourhood mobility in context: household moves and changing neighbourhoods in the Netherlands" *Environment and Planning A* **41** 1442-1459
- van Ham M, Feijten P, 2008, "Who wants to leave the neighbourhood? The effect of being different from the neighbourhood population on wishes to move" *Environment and Planning A* **40** 1151-1170
- van Ham M, Manley D, 2009, "Social Housing Allocation, Choice and Neighbourhood Ethnic Mix in England" *Journal of Housing and the Built Environment* **24** 407-422

van Ham M, Manley D, 2010, "The effect of neighbourhood housing tenure mix on labour market outcomes: a longitudinal investigation of neighbourhood effects"
Journal of Economic Geography **10** 257-282

Table 1. Descriptive statistics of neighbourhoods in 2001 (excludes neighbourhood with under 50 inhabitants)

	Mean	Std Dev.	Minimum	Maximum
Number of Inhabitants	1103	817.9	68	3251
% Foreign Born	13.5	9.4	2.8	51.0
% With Children under 18	44.3	18.1	8.1	86.1
% Students	14.2	10.0	4.9	64.3
% With Low Education (compulsory school)	12.3	5.3	1.1	26.4
% Non-employed (age 20-64)	19.5	9.2	4.1	47.6
% Receiving Social Benefits	2.9	3.9	0.0	23.1
% Public Renters	9.2	19.3	0.0	95.8
Median Annual Disposable Income (100000 SEK ^a)	2.7	1.0	0.8	4.8
N	122			
^a 10 SEK is about €1				

Source: Author's calculations using data from GeoSweden.

Table 2. Descriptives of moving households between 1997 and 2006, in comparison to all households in the city in 2001 (percentages except income).

	Moving households 1997-2006	All households 2001
Ethnicity		
All Swedish born	79.4	81.6
All Foreign born	17.2	15.2
Mixed ethnicity	3.4	3.2
Family Composition		
Couple	6.5	11.8
Couple with children	12.4	16.3
Single Parent	6.4	7.6
Single	74.7	64.4
Student in household	30.8	19
Education		
Compulsory School	7.5	9.4
High School (incl. students)	55.5	43.5
University	31.5	35.7
Employment		
All Employed	55.5	54.6
All Non-Employed	28.0	21.4
Mixed Employment status	3.7	4.6
Age 65+	7.5	19
Mean Disposable Income (100 000 SEK)	2.02	2.09
Social Allowances Recipient	8.2	4.7
Tenure		
Home Ownership	18.7	17.2
Cooperative Ownership	24.3	39.4
Private Rental	17.8	19.6
Public Rental	11.5	10
N	130,231	79,143

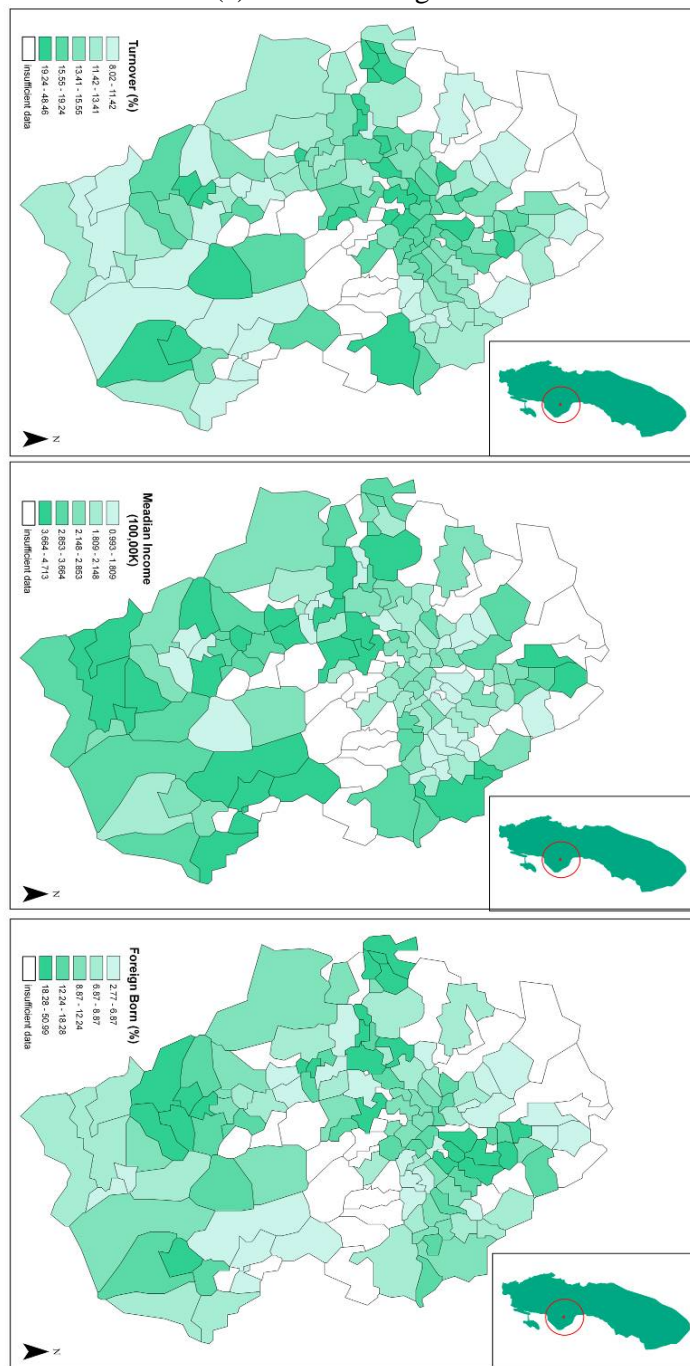
Source: Author's calculations using data from GeoSweden.

Table 3. Descriptive statistics of neighbourhoods in quintiles according to turnover rates (percentages except age and income).

	Quintiles of turnover rates, quintile 1 = lowest turnover rate									
	1		2		3		4		5	
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Turnover Rate	6.9	2.5	11.5	0.9	14.2	0.7	17.3	1.1	26.8	7.6
Number of Inhabitants	607.2	572.9	741.4	561.8	1394.7	629.2	1319.0	876.5	1458.8	996.9
Share Foreign Born	7.7	2.9	9.3	4.4	14.1	6.8	16.6	10.8	19.6	12.7
Share with Children	52.4	13.2	56.9	14.2	41.1	16.2	38.4	18.2	35.6	17.4
Share Students	9.0	1.9	10.3	2.2	11.7	2.3	14.1	2.0	25.9	17.6
Share with Low Education	12.8	4.7	12.7	3.9	14.5	5.0	12.8	5.2	8.5	6.0
Share Unemployed (20-64)	13.0	4.7	14.3	5.1	19.7	6.7	23.1	8.0	27.6	11.3
Median Neighbourhood Income (100 000 SEK)	3.7	0.7	3.3	0.8	2.4	0.7	2.1	0.5	1.8	0.7
Share Receivers of Social Allowances	0.5	0.9	1.4	1.8	3.5	2.9	4.5	4.5	4.6	5.5
Share Public Renters	1.1	3.4	7.5	20.6	13.6	21.6	14.7	23.8	8.7	18.1
NR of neighbourhoods	24		25		24		25		24	

Source: Author's calculations using data from GeoSweden.

Figure 1. Maps showing Uppsala neighbourhoods according to (a) turnover rate, (b) median income and (c) share of foreign born



Source: Author's calculations using data from GeoSweden.

Table 4. Conditional Logit Model of neighbourhood choice

	Model 1			Model 2		
	Coefficient	Std. Error	Sign.	Coefficient	Std. Error	Sign.
Median Neighbourhood Income (100 000 SEK) (MI)						
MI x Income in 1st Quintile (lowest)	-1.3168	0.0075	***	-1.0398	0.0110	***
MI x Income in 2nd Quintile	-0.9085	0.0077	***	-0.9580	0.0109	***
MI x Income in 3rd Quintile	-0.6300	0.0086	***	-0.7724	0.0114	***
MI x Income in 4th Quintile	-0.3143	0.0096	***	-0.5541	0.0124	***
MI x Income in 5th Quintile	-0.0194	0.0105		-0.3047	0.0131	***
Share Foreign born (FB)						
FB x Swedish born				-0.0082	0.0007	***
FB x Foreign born				0.0328	0.0008	***
FB x Mixed Ethnicity Household				0.0069	0.0018	***
Share with Children (Ch)						
Ch x Couple				0.0008	0.0007	
Ch x Couple with Children				0.0217	0.0006	***
Ch x Single Parent				-0.0001	0.0005	
Ch x Single				-0.0112	0.0004	***
Share Students (St)						
St x Student in Household				0.0335	0.0004	***
St x No Student in Household				-0.0199	0.0005	***
Share with Low Education (LE)						
LE x No High School Degree				0.0871	0.0025	***
LE x High School Degree				0.0066	0.0009	***
LE x University Education				-0.0495	0.0013	***
Share Non-Employed (20-64) (NE)						
NE x All Employed				-0.0363	0.0007	***
NE x All Non-employed (20-64)				-0.0247	0.0007	***
NE x Mixed Employment status (20-64)				-0.0272	0.0015	***
NE x Retired (age 65 or older)				-0.0402	0.0014	***
Share Social Allowance Recipients (SA)						
SW x Household receive Social Allowances				0.1175	0.0019	***
SW x Household do not receive Social Allowances				0.0624	0.0013	***
Share Public Renters (PuR)						
PuR x Home Owner				-0.0064	0.0004	***
PuR x Owner of Cooperative Flat				-0.0052	0.0003	***
PuR x Private Renter				-0.0042	0.0004	***
PuR x Public Renter				0.0095	0.0004	***
Observations	1302310			1302310		
Log Likelihood	-263067.47			-245947.17		
Pseudo-R ²	0.1263			0.1831		

Source: Author's calculations using data from GeoSweden.