Applied Spatial Analysis and Policy
https://doi.org/10.1007/s12061-021-09411-5

Modelling Neighbourhood Effects in Three Dutch Cities Controlling for Selection

Agata A. Troost[1] · Maarten van Ham[1,2] · Heleen J. Janssen[3]

Received: 20 November 2020 / Accepted: 30 June 2021

Abstract

The non-random selection of people into neighbourhoods complicates the estimation of causal neighbourhood effects on individual outcomes. Measured neighbourhood effects could be the result of characteristics of the neighbourhood context, but they could also result from people selecting into neighbourhoods based on their preferences, income, and the availability of alternative housing. This paper examines how the neighbourhood effect on individual income is altered when geographic selection correction terms are added as controls, and how these results vary across three Dutch urban regions. We use a two-step approach in which we first model neighbourhood selection, and then include neighbourhood choice correction components in a model estimating neighbourhood effects on individual income. Using longitudinal register datasets for three major Dutch cities: Amsterdam, Utrecht and Rotterdam, and multilevel models, we analysed the effects for individuals who moved during a 5-year period. We show that in all cities, the effect of average neighbourhood income on individual income becomes much smaller after controlling for explicitly modelled neighbourhood selection. This suggests that studies that do not control for neighbourhood selection most likely overestimate the size of neighbourhood effects. For all models, the effects of neighbourhood income are strongest in Rotterdam, followed by Amsterdam and Utrecht.

Keywords Neighbourhood effects · Neighbourhood selection · Selection bias · Income · Social inequality

✉ Agata A. Troost
a.a.troost@tudelft.nl

1 Department of Urbanism, Faculty of Architecture and the Built Environment, Delft University of Technology, P.O. Box 5043, 2600 GA Delft, The Netherlands

2 School of Geography and Sustainable Development, University of St Andrews, Scotland, UK

3 Independent Research Group Space, Contexts and Crime Max Planck Institute for the Study of Crime, Security and Law, Freiburg, Germany

Published online: 24 August 2021

Introduction

A major challenge in the field of neighbourhood effects is the estimation of contextual effects free of bias from the non-random selection of households into neighbourhoods. Neighbourhood effects are causal effects of the spatial context on individual outcomes, such as education, health or income (Galster & Sharkey, 2017). However, people do not randomly choose their residential neighbourhood, but generally select neighbourhoods inhabited by people with similar characteristics to themselves (Hedman et al., 2011). The observed "effect" of neighbourhood characteristics on individual outcomes is therefore at least partly the result of selective sorting of individuals into neighbourhoods (Hedman & Van Ham, 2012).

The concern about selection bias is not new. Sampson (2008) traced it back to the early work of Jencks and Mayer (1990), who concluded that we do not know whether differences in outcomes result from neighbourhood factors, rather than just from selection into the neighbourhoods. Researchers developed several strategies to overcome selection bias in their research designs and statistical modelling approaches, such as: experimental, or quasi-experimental research designs (Katz et al., 2001; Leventhal & Brooks-Gunn, 2003; Sanbonmatsu et al., 2006); the use of instrumental variables (Hedman & Galster, 2013) or sibling data (Oreopoulos, 2003; Zick et al., 2013); or longitudinal data and fixed effects models (Boone-Heinonen et al., 2011; Jokela, 2015). Another approach explicitly models neighbourhood selection as a conditional logit regression, in which the probability of an individual choosing a neighbourhood based on individual, household and neighbourhood characteristics is estimated (Ioannides & Zabel, 2008). In a second step these predicted probabilities are turned into correction components, which are used as controls in a neighbourhood effects model for the same subjects. This method was applied and refined in a study of neighbourhood effects on individual income by Van Ham et al. (2018).

Because this two-step approach explicitly models neighbourhood selection, it can be used to provide insights about the influence of local selection mechanisms on neighbourhood effects, when it is applied to multiple cities. Small and Feldman (2012) argued that differences in neighbourhood effects between cities are often overlooked in neighbourhood effects studies. The current study contributes to this discussion by arguing that dealing with selection bias can have varying results in different cities. We examine selection mechanisms and their influence on neighbourhood effects in three different Dutch urban regions (which from now on will be also referred to as cities): Utrecht, Amsterdam and Rotterdam. Using longitudinal and geo-coded register data, we first model neighbourhood selection, and then the effect of average neighbourhood income on individual income from work. We further develop the original two-step approach in several ways. Firstly, where previous studies only included individuals who moved within one particular year, which biases the sample towards "frequent movers" (Hansen & Gottschalk, 2006; Phinney, 2013), we included individuals who moved within a period of 5 years (2010 to 2014). Secondly, it is expected that moving to a higher income neighbourhood would lead to an *increase* in individual earned income. We therefore investigate neighbourhood effects on both income in a given year, and change in income. Finally, whereas Van Ham et al. (2018) used clustered standard errors to control for the clustering of individuals in neighbourhoods, we use a multilevel model, which is a more appropriate choice for modelling clustered data (Jones, 1991; Cheah, 2009), and gives us insight into the amount of variance explained by the predictors on neighbourhood level.

Theoretical Background

Selection Bias in the Neighbourhood Effects Studies

Concerns about selection bias, or neighbourhood effects being overestimated due to the non-random structured selection of people into neighbourhoods, have long plagued the field of neighbourhood effects (Small & Feldman, 2012). In response to these concerns, different approaches were taken to take selection bias into account. Quasi-experimental designs were developed, such as the Moving to Opportunity scheme, beginning in 1994, in which low income American families could move out from high-poverty public housing to low-poverty middle class neighbourhoods (Leventhal & Brooks-Gunn, 2003). Oreopoulos (2003) took another approach by using data on public housing inhabitants from Toronto; he argued that they do not select their neighbourhoods themselves, which should minimise selection bias. Boone-Heinonen et al. (2011) noted that ideally observational designs should be longitudinal, which allows for assessing changes in individual's characteristics in relation to changes in the neighbourhood. Cross-sectional studies often control for variables regarded as proxies for selection bias, and instrumental variables related to the proposed predictors, but not the outcomes (Zick et al., 2013). Such strategies, however, give little insight in the selection process itself, and in the case of self-reported preferences might even introduce additional bias (Boone-Heinonen et al., 2011).

But could there be *too much* focus on selection? Sampson (2012) observes that the function of individual choice remains a controversial matter: some researchers argue that "segregation and constraints of inequality override choice, in extreme case almost as if individuals are pawns in a predetermined game"; others "valorise choice to the point where it is said to undercut research efforts to investigate the effects of neighbourhood context" (p. 287). Such discussions hark back to the debate on the influence of socioeconomic structures vs individual agency on one's life outcomes. A few studies attempt to have the best of both worlds by modelling the effect of individual moving choices on the socioeconomic structure of neighbourhoods (Sampson & Sharkey, 2008; Hedman et al., 2011). Hedman et al. (2011) demonstrate that people move to neighbourhoods with inhabitants of income, ethnicity and family composition similar to their own, most often reproducing existing neighbourhood structures.

Ioannides and Zabel (2008) recognised the importance of explicitly modelling neighbourhood selection in their study of housing demand. They use a conditional logit model of selection from a set of 11 tracts (ten of them randomly selected from all US census tracts in the metropolitan area, plus the chosen tract of an individual). Subsequently, they deal with selection bias by deriving neighbourhood selection correction components from the first step, and including them in the models of housing demand. The results show that the neighbourhood effects became stronger after controlling for neighbourhood selection. In contrast, the results by Van Ham et al. (2018), who followed a very similar two-step strategy, show that the observed neighbourhood effects on individual income weaken after adding the neighbourhood selection controls. The different effects could be explained by the different dependent variables used in the two studies, as well as differences in their measurement of the neighbourhood context and different residential preferences of the American and Dutch households. Another explanation might be that Van Ham and colleagues used the full set of neighbourhoods, rather than a random subset like Ioannides and Zabel did. In any case, Ioannides and Zabel's paper introduces a convincing method of combining both the neighbourhood selection and effects processes in one modelling approach.

Neighbourhood Selection

Over the life course people choose dwellings in different locations and neighbourhood types, each suited to their current preferences and resources, and closely related to major life events such as starting a family or retiring (Rossi, 1955). Because dwelling types are not randomly distributed over urban space, households tend to move to neighbourhoods with households similar to themselves, since they prefer to live in similar housing.

These trends can change over time; throughout the second half of the 20th century, Western middle class people used to move to the suburbs to raise their children in big dwellings and safe, sleepy neighbourhoods. Recently, more and more higher-educated couples refuse to forsake the inner city services and institutions once they have children (Boterman, 2012). Especially the mothers in dual-earner households benefit from the easy access to city resources (Boterman & Bridge, 2015). The demand for inner city dwellings rises, spreading from the very city centre to the surrounding, previously undesirable neighbourhoods. Access to workplaces can also affect neighbourhood choice. Depending on the chosen mode of transportation, neighbourhoods located close to the highway or train station—or both—might be more attractive (Van Ham et al., 2001).

In their literature review on neighbourhood selection, Hedman et al. (2011) list empirical evidence confirming the position of income and ethnicity as the main drivers of neighbourhood choice and resulting patterns of segregation (see also Musterd et al., 2016; Galster & Turner, 2019). The choice-limiting effect of income is straightforward: low income households cannot afford to live in expensive neighbourhoods. Such "affordability constraints" can be accounted for by including dwelling value and household economic resources in the model (Bruch & Mare, 2012).

The effect of ethnicity is more complicated, also because ethnicity and income are strongly related. Many researchers, starting from the famous models by Schelling

(1971), emphasise the importance of individual preferences with regard to the ethnicity of their neighbours in understanding neighbourhood choice, and the resulting patterns of segregation. While most minorities earn less than the ethnic majority, many of them have other economic, religious and cultural reasons for living close to each other (McAvay, 2018). Whereas an ethnic majority family most likely will move out of a poverty neighbourhood when they can afford it, an ethnic minority family which could live in a more affluent, overwhelmingly native neighbourhood, might choose to remain in a neighbourhood with people from their own background in order to stay in touch with family and cultural traditions (Boschman & Van Ham, 2015).

Next to preferences and affordability constraints, selection can be shaped by a limited access to information about neighbourhoods. Although such information is relatively easy to access in the Netherlands, movers might not consider the parts of the city they do not know well, and such knowledge is often closely related to the income- and ethnicity-based segregation (Krysan & Crowder, 2017). People's social networks are related to the places where they already live, and if these networks comprise, for example, high income individuals, it is likely that members of the network will share information mostly about higher income neighbourhoods.

Different Cities, Different Neighbourhood Processes

Studies of neighbourhood effects often overlook that both selection and neighbourhood effects can vary depending on local circumstances. For a long time, Chicago was seen as the prototype city to study neighbourhood effects (Small & Feldman, 2012). Key authors such as Wilson (1987) have claimed that various characteristics observed in Chicago, such as low density of local businesses and institutions in poor neighbourhoods, are representative for all US cities. However, Small and Feldman (2012) show that local establishment density in poor Chicago neighbourhoods is much lower than the averages for other American cities. Furthermore, Burdick-Will et al. (2011) found significant effects of the MTO experiment on children test scores in Chicago and Baltimore, but not in Los Angeles, Boston and New York. Because of such findings, Small and Feldman (2012) call for more neighbourhood effects studies with data collected from "average", middle-sized cities without unusually high crime or poverty levels. In Europe, Musterd et al. (2012) also found that evidence of neighbourhood income mix effects on individual income varies between Swedish cities.

Because of the need for more heterogeneous research settings, and because of the importance of regional housing markets in understanding processes of neighbourhood selection, the current study focusses on three different urban regions in the Netherlands: Utrecht, Amsterdam and Rotterdam. Although all three cities are a part of the Randstad,[1] this metropolitan area "has no institutional foundation and

[1] A polycentric metropolitan area in the western part of the Netherlands, comprising the four largest Dutch cities (Amsterdam, Rotterdam, The Hague and Utrecht), as well as multiple smaller cities and the less-densely populated, agricultural core (known as the Green Heart), totalling a population of around 7 million (Stead & Meijers, 2015).

no formal powers of decision-making" (Stead & Meijers, 2015, p. 4), which leads to differences in local policies. Because of this and other, historical, reasons the three cities have developed differently. Utrecht has a far lower percentage of ethnic minorities (32%) compared to Amsterdam (50%) and Rotterdam (49%).[2] Rotterdam has a higher share of households with a lower income and lower educational level compared to the other cities. This situation is related to the city's industrial past: even though now it has a university like Utrecht and Amsterdam (which has two), much of its labour market revolves around its port, the largest in Europe (Stead & Meijers, 2015). Amsterdam is a leading cultural and financial centre with a large number of both high- and low-income immigrants, as well as Dutch citizens with foreign roots. Utrecht has a similar labour market to Amsterdam, with an overrepresentation of high socio-economic status occupations. In Utrecht, ethnic minorities are concentrated in suburban districts, where most of the city's social housing is located; the overall percentage of social housing is also the lowest of the three cities (33% in 2015, compared to 45% in Rotterdam and 43% in Amsterdam[3]). Unlike in Utrecht, social housing is quite evenly spatially distributed in Amsterdam and Rotterdam. In Amsterdam this distribution is an effect of decades of housing policy which prioritised social mix on the neighbourhood level. Rotterdam also has a high percentage of social housing throughout the city, even after large scale urban restructuring since the 1990s.

These differences in local economies between the three cities are likely to affect the process through which households select into different types of neighbourhoods. Amsterdam and Utrecht are both known for their very high housing prices, and in these cities the middle income households are in a difficult situation, as they earn too much to qualify for social housing, but cannot afford to live in the most desirable neighbourhoods. As a result, income might be a stronger predictor of neighbourhood selection in Utrecht and Amsterdam compared to Rotterdam. Other personal characteristics, such as education and family composition, combined with local particularities, may also lead to differences in selection patterns between the cities, as well as to differences in the magnitude of neighbourhood effects.

Neighbourhood Effects on Income and Income Change

Many empirical studies have investigated neighbourhood effects on individual income (see Galster & Sharkey, 2017, p. 21, for an overview). The evidence suggests that the neighbourhood influence in adulthood is weaker than the influence experienced in childhood, when social networks are more often limited to the neighbourhood and major career choices are yet to be made (Galster & Sharkey, 2017). Still, adults can become more similar to their neighbours through the social-interactive mechanisms of adapting to their behaviours, aspirations and attitudes, conforming to local social norms or accessing information and resources through their social

[2] https://opendata.cbs.nl/statline/

[3] https://opendata.cbs.nl/statline/

networks (Galster, 2012). With regards to income, someone can find a better paid job through the neighbourhood social network. On the other hand, a person living in an area with a high level of unemployment and many low-skilled employees will more readily accept her own low salary (Sari, 2012). Pinkster (2007, 2014) investigated these mechanisms in her qualitative studies of a high-poverty neighbourhood in the Hague, Transvaal-Noord. She found that the local dense social networks did help with finding a low-skilled job, but the availability of such jobs discouraged the immigrant neighbourhood inhabitants from pursuing further education, learning Dutch and familiarising themselves with the formal job market. This illustrates how the information and contacts in the neighbourhood can influence individuals' income.

For employed household heads moving to a new neighbourhood, a change in their income can be caused primarily by the aforementioned social mechanisms, but also by a number of other neighbourhood characteristics. In their study of the influence of neighbourhood social mix on Stockholm's inhabitants' income, Galster et al. (2016) describe the potential influence of local crime levels, institutional resources and job accessibility, widely studied in the literature. Neighbourhood average income tends to be a good proxy for these characteristics because of local self-reinforcing processes of spatial inequality—for example, well-connected neighbourhoods tend to have more expensive dwellings, which attracts richer inhabitants, who further profit from the easy access to jobs and services (Toft & Ljunggren, 2016). Such processes correspond to theories of the Matthew effect (Merton, 1968) and cumulative disadvantage (Sampson & Laub, 1997). A move to an affluent neighbourhood should therefore expose an individual to many beneficial resources, regardless of her previous socioeconomic status. It can also inspire an aspirational attitude of trying to „keep up" with the richer neighbours, encountered at local establishments and events.

Current Study

Following Ioannides and Zabel (2008) and Van Ham et al. (2018), we first explicitly model neighbourhood selection and then use the predicted probabilities from this model to construct correction components, which we use as control for neighbourhood selection in the second modelling step: the neighbourhood effects models. We hypothesise that the correction components, which control for individual characteristics as well as neighbourhood selection, will reduce the selection bias present in the observed neighbourhood effect more accurately than the individual characteristics variables. In other words, we expect that the influence of neighbourhood income on individual income becomes smaller after including personal characteristics, and further diminishes after controlling for selection bias by using correction components.

Crucially, we model neighbourhood selection and effects with data from Utrecht, Amsterdam and Rotterdam. We expect that the degree to which the effects in the final models are affected by controlling for neighbourhood selection will be different in the three cities. That is because the neighbourhood selection

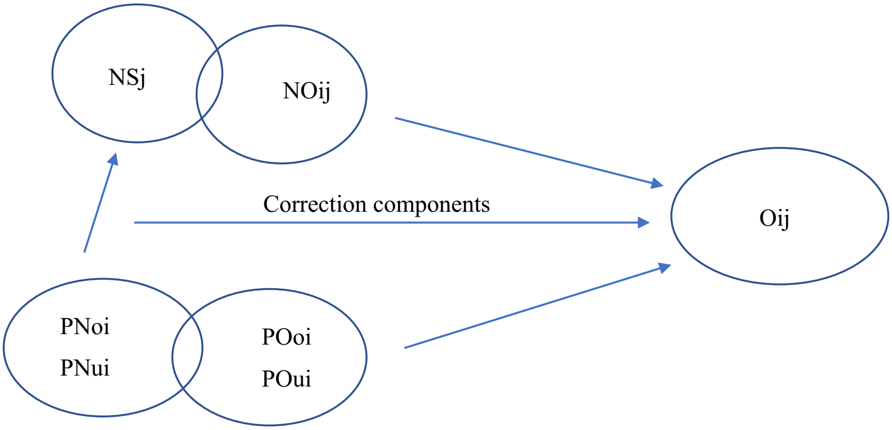

Fig. 1 Relationships between the observed and unobserved variables on neighbourhood and individual level

controls reflect important differences between local housing markets, which can limit individuals' choice to various degree, based on e.g. the availability of social housing or family-sized dwellings. However, there can be also other unobserved neighbourhood-related, city-level factors which explain the differences between the three cities; for example, effects from differing regional policies affecting the local economy.

To make the relationships between the observed and unobserved factors in the models clearer, we have drawn a diagram showing these relationships (Fig. 1). In the selection models, individuals (i) have the observed and unobserved personal characteristics affecting neighbourhood selection (PNoi, PNui), which relate to the characteristics of the neighbourhoods (j) they choose (NSj). In the neighbourhood effects models, the chosen neighbourhood can then have an influence on the final outcome, individual income (Oij), with their characteristics (NOij), which might not be entirely accounted for by the variables in the selection models. At the same time, individuals have other characteristics which affect Oij, also observed and unobserved (POoi and POui), in the neighbourhood effects model. Some of them (such as previous income or education) can overlap with the observed and unobserved personal characteristics affecting selection (PNoi and PNui), but not all of them will and those that do might not overlap entirely. Similarly, the neighbourhood characteristics (NOij) which influence individual income (Oij) are not necessarily the same which have led to choosing that neighbourhood (NSj). In our approach, we account for the bias caused by the statistical relationship between personal characteristics and neighbourhood selection criteria (PNi and NSj) by including the correction components based on the selection models in the neighbourhood effects models. Because of possible differences between the (especially unobserved) personal characteristics affecting selection (PNi) and individual income (POi), as well as neighbourhood characteristics relevant to selection (NSj) and individual income (NOij), which can be also subjected to city-wide trends, we cannot say that these correction components cover all the city-level differences. Still, they should reflect many of these differences because of the wide range of variables used in both sets of models.

Data and Methods

Sample and Data

We used data from the Netherlands Social Statistical Database (SSD), a population registry including individual level, geo-coded longitudinal data for the entire population of the Netherlands from 1999 onwards. These data are merged with register-based publicly available neighbourhood-level data from Statistics Netherlands. We selected household heads with income from work in 2015 who moved between 2010 and 2014. This resulted in a final dataset containing 54,045 individuals in Utrecht, 84,935 in Amsterdam and 59,681 in Rotterdam, which corresponds to the different population sizes of these cities. We included both people moving within the city, and people moving in from other parts of the Netherlands. We only included household heads in our sample because including multiple earners from the same family would complicate the data structure, and second earners in the Netherlands, especially mothers, very often work part-time (Endendijk et al., 2018), which is not accounted for in the available income variables. Still, the sample consists of 40% women in Amsterdam and around 37% in both Rotterdam and Utrecht.

We included data of movers from multiple years to increase the external validity of our study. Only the last move is included, regardless of whether someone moved a year or 5 years before; we then control for time in our models. The previous studies have only used moves from 1 year (Ioannides & Zabel, 2008; Van Ham et al., 2018), which can lead to an overrepresentation of recent movers, such as foreign-born, single and young people (Hansen & Gottschalk, 2006; Hedman et al., 2011). People who moved outside of the studied cities are not in the dataset, whereas people who moved into the cities are included. The spatial units we use are the municipalities' administrative neighbourhoods, which are likely to resemble neighbourhoods as people know them. Permentier et al., (2008, p. 840) conducted a survey in which the inhabitants of Utrecht asked to name their neighbourhood either gave the same (81.5%) or a similar (14%) name to the official name of the administrative area. This suggests that, in Dutch cities, people identify their neighbourhood similarly to the official administrative names (although the spatial bounds they have in mind might not map exactly onto the administrative borders). On average, there are 144 individuals from our sample per neighbourhood in the Utrecht region, 298 in Amsterdam and 195 in Rotterdam (for the total population, the average is 1812 individuals per neighbourhood in Utrecht, 1403 in Amsterdam and 1331 in Rotterdam). We used the urban regions of Utrecht, Amsterdam and Rotterdam, which include the main city, but also surrounding municipalities, representing regional housing markets (see the appendix for a list of included municipalities).

Analytical Approach and Variables

Our modelling approach has two steps. It is akin to the Heckman 2-step solution overcoming selection bias (Mroz, 1987; Winship & Mare, 1992). We first estimate neighbourhood selection with a conditional logit model, in which we model the probability that an individual chooses a particular neighbourhood from the complete choice set of all neighbourhoods in the city (i.e. Utrecht, Amsterdam or Rotterdam) and its surrounding suburbs. Because of the properties of the conditional logit model, individual characteristics can only be included as interactions with neighbourhood characteristics. Based on these interactions we estimate linear probabilities of an individual choosing each neighbourhood in the choice set. These, in turn, are converted into correction terms by using a technique similar to the Inverse Mills Ratios, following Van Ham et al. (2018) and Ioannides and Zabel (2008). Because many neighbourhoods are similar on particular characteristics, and people tend to choose their neighbourhoods in a very structured way, we observe high levels of collinearity between the correction terms. Therefore, instead of keeping hundreds of correction terms, we reduce them to a smaller set of correction terms using Principal Component Analysis (for a more detailed description, see Van Ham et al., 2018). These components reflect the probabilities that types of households select types of neighbourhoods.

In the selection model we included neighbourhood-level variables (average dwelling value, number of restaurants within 3 km, distance to train station, distance to highway access lane, share of dwellings built after 2000, share of non-Western minorities, share of social housing, share of private rental, share of single person households and share of households with children), as well as individual-level variables (age, ethnic background, family composition, household income and education level). We included these characteristics for the year when the individual moved (2010–2014), except for educational level for which we included obtained educational level in 2015. Education level is measured by four dummy variables indicating "lower educated" for people with unfinished secondary education, "middle educated" for those with a secondary or practical vocational (*mbo*) degree and "higher educated" for people with a bachelor or higher degree (*wo* or *hbo*), and "missing" for those with missing information on their educational level (25%). Migration background is represented in the model by two dummy variables: "Western migrant origin" and "non-Western migrant origin", which identify individuals with at least one parent born abroad. "Western countries", according to the Statistics Netherlands definition, include all European and Northern American countries plus Japan, Australia and Indonesia.

In the second modelling step we model neighbourhood effects. We regressed individual earned income in 2015 (log-transformed) on neighbourhood and individual characteristics by estimating a multilevel model with individuals at level 1 nested in neighbourhoods at level 2. In Model 1 we included average neighbourhood income as a predictor at neighbourhood level. In Model 2 we added individual characteristics to the model. Model 3 includes the variables from the previous models plus the neighbourhood types correction components, which reflect the possibility of each person selecting a particular type of neighbourhood. Consequently, in Model

3 we directly control for neighbourhood selection. The differences in neighbourhood effects between Model 2 and Model 3 indicate to what extent neighbourhood effects might be overestimated when not explicitly controlling for selection. The importance of using a multilevel model for our data is confirmed by the results of the null-models (in the appendix), showing the amount of variance in individual income on individual and neighbourhood level. The intraclass correlation indicates that approximately 19% of the variance in individual income in Utrecht, 15% in Amsterdam and 18% in Rotterdam is on the neighbourhood level. Therefore, it is important to use multilevel modelling in order to correctly estimate standard errors. Furthermore, we can observe if and how unexplained variance on the neighbourhood level diminishes after including new predictors in the model.

In the neighbourhood effects models, the neighbourhood level variable is average neighbourhood income. The same individual characteristics as in the selection models are included, but with values taken from the 2015 datasets and with age in years instead of three age categories. We also included the number of months an individual has been living in the neighbourhood, to control for variation in exposure to neighbourhood conditions. The dependent variable is the logarithm of income from work in 2015. In order to directly model income change, in an additional set of models we included the logarithm of income in the year the individual moved to the neighbourhood. Furthermore, to check whether the differences in the neighbourhood average income coefficient size between the different cities are significant, we used formal tests for comparing the coefficients across models using different samples, described in detail by Paternoster et al. (1998).[4]

Results

Selection Models

The results from the first step of our approach, the conditional logit model (in the appendix), show the effects of the interactions between individual and neighbourhood level characteristics in predicting neighbourhood choice. Most of the coefficients are significant, revealing the structured selection of types of individuals into types of neighbourhoods. For example, couples with children are more likely to select neighbourhoods with already a high share of such household type. The results also indicate that ethnic minorities are more likely to move to neighbourhoods with a high share of people with a non-Western background. This effect is strongest in Amsterdam, followed by Rotterdam and then Utrecht. Whereas higher educated individuals in Utrecht and Amsterdam are more likely to select a neighbourhood with a larger share of ethnic minorities, in Rotterdam they are less likely to select this type of neighbourhood. Higher educated individuals are less likely to select a neighbourhood with a large share of social housing in Utrecht, but more likely in

[4] $Z = \frac{b_1 - b_2}{\sqrt{SE_{b_1}^2 + SE_{b_2}^2}}$

Amsterdam, and in Rotterdam we do not find an effect. Individuals with a non-Western migration background are less likely to select a neighbourhood with a high share of social housing in Amsterdam and Rotterdam, but not in Utrecht. Higher educated people select neighbourhoods with a high number of restaurants, a high share of singles, and a shorter distance to the train station, thus likely located in the city centre. Age differences between moving patterns in the three cities also emerge: for example, people below the age of 25 are more likely than 25–65 year old people to choose a neighbourhood with a high share of households with children in Rotterdam, but less likely to do so in Amsterdam, and the interaction is non-significant in Utrecht. These results of the selection model show that there are important differences in the determinants of neighbourhood choice between the three cities. This subsequently translates into the correction terms, which will be included in the neighbourhood effects model in the next step. Furthermore, the model fit differs between the three cities, with the Pseudo R-squared of 0.079 for Utrecht, 0.07 for Rotterdam and the lowest 0.057 for Amsterdam. These statistics suggest that the extent to which our model explains the selection is the biggest in Utrecht, even though one has to be cautious while comparing the Pseudo R-squared of models ran on different datasets (Tables 1, 2).

Effect Models: Income

Tables 3, 4, 5 present the results of the multilevel models predicting individual income from work. In Model 1 we find a positive statistically significant effect of neighbourhood income for all cities, meaning that the higher the neighbourhood average income, the higher individual income a few years after the move. However, the magnitude of the effect differs between the cities, with a significantly stronger neighbourhood effect in Rotterdam compared to Utrecht and Amsterdam.[5] This contradicts our expectation that there are no differences in the strength of neighbourhood effects between the three cities. A €10,000 difference in average neighbourhood income is related to a 24% difference in individual income in Utrecht, a 27% difference in Amsterdam, and a 32% difference in Rotterdam.

In Model 2 we included individual characteristics in addition to the average neighbourhood income. In all three cities the effect of average neighbourhood income on individual income drops in size. In Utrecht the effect of average neighbourhood income drops by 42%, in Amsterdam by 33% and in Rotterdam by 44%. The effects of individual characteristics are similar across all three cities: all effects are statistically significant ($p < 0.001$) and show that ethnic minority members have a lower income, while couples, older and higher educated individuals have a higher income. The residual variances at neighbourhood and individual level are significantly reduced in Model 2 compared to Model 1, indicating that the individual

[5] The effect of neighbourhood average income in Model 1 is significantly larger in Rotterdam compared to Amsterdam ($Z = \frac{.238-.279}{\sqrt{.007^2+.009^2}} = -3.60$). The effect is not significantly different between Amsterdam and Utrecht ($Z = \frac{.238-.218}{\sqrt{.007^2+.008^2}} = 1.88$).

Table 1 Descriptive statistics of the neighbourhood and individual level variables used in the selection models

	Utrecht				Amsterdam				Rotterdam			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Neighbourhood characteristics												
Average dwelling value (in 1000 Euros)	252.44	96.77	*	*	243.9	93.27	*	*	168.95	65.82	*	*
Restaurants within 3 km	81.42	99.36	0	638.8	317.37	383.78	0	1259.1	106.25	120.11	1.9	405.5
Distance to train station (km)	2.88	2.55	0.3	13.1	2.79	1.89	0.4	9.4	2.85	2.15	0.2	10.9
Distance to highway access lane (km)	1.86	0.76	0.1	6.6	1.85	0.90	0.2	4.4	1.92	.86	0.1	4.9
Share of dwellings built > 2000	17.45	29.71	0	100	14.41	22.27	0	100	14.66	23.12	0	100
Share of non-Western minorities	15.76	13.07	0	78	28.15	18.69	0	78	30.40	18.67	0	80
Share of social housing	31.01	21.47	0	100	40.63	20.86	0	100	41.45	21.56	0	100
Share of private rental	15.23	10.88	0	94	23.62	16.04	0	100	17.35	12.48	0	100
Share of singles	41.37	16.1	10	96	49.14	13.67	8	94	43.98	12.74	0	93
Share of households with children	32.55	13.04	0	72	28.49	11.58	1	68	30.98	10.73	0	70
Individual characteristics												
Age < 25	0.08		0	1	0.07		0	1	0.11		0	1
25–65	0.92		0	1	0.93		0	1	0.89		0	1
> 65	0.0004		0	1	0.0005		0	1	0.0006		0	1
Ethnicity Native Dutch	0.77		0	1	0.56		0	1	0.57		0	1
Western migrant origin	0.11		0	1	0.20		0	1	0.15		0	1
non-Western migrant origin	0.12		0	1	0.24		0	1	0.28		0	1
Family type Single	0.38		0	1	0.48		0	1	0.43		0	1
Couple	0.29		0	1	0.26		0	1	0.25		0	1
Couple with children	0.25		0	1	0.17		0	1	0.21		0	1
Other family type	0.08		0	1	0.09		0	1	0.11		0	1
Gross household income (in 10,000 euro)	2.56	1.22	*	*	2.56	1.8	*	*	2.31	1.19	*	*
Educational level Low	0.07		0	1	0.08		0	1	0.12		0	1

Table 1 (continued)

	Utrecht				Amsterdam				Rotterdam			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Middle	0.22		0	1	0.20		0	1	0.30		0	1
Higher	0.48		0	1	0.46		0	1	0.34		0	1
Education missing	0.23		0	1	0.26		0	1	0.24		0	1

Notes: *We are not able to show minimum and maximum due to Statistics Netherlands disclosure restrictions. In the Utrecht urban region, we analysed 54,045 individuals in 375 neighbourhoods; in Amsterdam, 84,935 in 285; and in Rotterdam, 59,681 in 306

Table 2 Descriptive statistics of the neighbourhood and individual level variables used in the effect models (2015)

	Utrecht				Amsterdam				Rotterdam			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Log income in 2015	10.71	0.55	*	*	10.71	.67	*	*	10.58	0.59	*	*
Average neighbourhood income (in 10,000 euro)	3.45	0.83	*	*	3.44	1	*	*	2.99	0.75	*	*
Log income in the year of move	10.62	0.49	*	*	10.62	.6	*	*	10.5	0.52	*	*
Native Dutch	0.77		0	1	0.55		0	1	0.57		0	1
Non-Western migrant origin	0.12		0	1	0.24		0	1	0.28		0	1
Western migrant origin	0.11		0	1	0.20		0	1	0.15		0	1
Single	0.38		0	1	0.48		0	1	0.43		0	1
Couple	0.26		0	1	0.26		0	1	0.24		0	1
Couple with children	0.34		0	1	0.25		0	1	0.29		0	1
Other family type	0.09		0	1	0.10		0	1	0.12		0	1
Age	38.39	10.29	18	85	37.15	9.31	18	90	37.7	10.1	18	85
Lower educated	0.07		0	1	0.08		0	1	0.12		0	1
Middle educated	0.22		0	1	0.20		0	1	0.30		0	1
Higher educated	0.48		0	1	0.46		0	1	0.34		0	1
Education missing	0.23		0	12	0.26		0	1	0.24		0	1
Female	0.36		0	1	0.40		0	1	0.37		0	1
Months since move	39.67	17.82	12	71	36.4	17.67	12	71	38.78	18.24	12	71

*We are not able to show minimum and maximum due to Statistics Netherlands disclosure restrictions

Table 3 Neighbourhood effects on income: Utrecht

	Model 1			Model 2			Model 3		
	B	SE	Exp	B	SE	Exp	B	SE	Exp
Average neighbourhood income (in 10,000 euro)	0.218***	(0.008)	1.244	0.131***	(0.004)	1.140	0.061***	(0.003)	1.063
Non-Western minority				− 0.117***	(0.006)	0.890	1.359***	(0.048)	3.892
Western minority				− 0.071***	(0.006)	0.931	0.269***	(0.013)	1.309
Couple				0.108***	(0.005)	1.114	0.054***	(0.006)	1.055
Couple with children				0.158***	(0.005)	1.171	− 0.003	(0.007)	0.997
Other family type				− 0.061***	(0.008)	0.941	0.101***	(0.009)	1.106
Age				0.057***	(0.002)	1.059	0.010***	(0.002)	1.010
Age squared				− 0.001***	(0.000)	0.999	− 0.000***	(0.000)	1.000
Middle educated				0.237***	(0.008)	1.267	1.180***	(0.021)	3.254
Higher educated				0.584***	(0.008)	1.793	0.756***	(0.030)	2.130
Education missing				0.342***	(0.009)	1.408	0.932***	(0.022)	2.540
Female				− 0.195***	(0.004)	0.823	− 0.136***	(0.004)	0.873
Months since move				0.001***	(0.000)	1.001	− 0.001*	(0.000)	0.999
Correction components included	NO			NO			YES		
Constant	9.952***	(0.028)		8.496***	(0.036)		9.287***	(0.042)	
Residual variance at neighbourhood level	0.013***	(0.001)		0.003***	(0.000)		0.001***	(0.000)	
Residual variance at individual level	0.259***	(0.001)		0.196***	(0.001)		0.148***	(0.000)	
N	54,045			54,045			54,045		

Standard errors in parentheses

*$p < 0.05$

**$p < 0.01$

***$p < 0.001$

Table 4 Neighbourhood effects on income: Amsterdam

	Model 1			Model 2			Model 3		
	B	SE	Exp	B	SE	Exp	B	SE	Exp
Average neighbourhood income (in 10,000 euro)	0.238***	(0.007)	1.269	0.166***	(0.005)	1.181	0.079***	(0.004)	1.082
Non-Western minority				– 0.142***	(0.005)	0.868	1.796***	(0.035)	6.025
Western minority				– 0.032***	(0.005)	0.969	0.315***	(0.012)	1.370
Couple				0.127***	(0.005)	1.135	0.049***	(0.005)	1.050
Couple with children				0.188***	(0.005)	1.207	0.058***	(0.006)	1.060
Other family type				– 0.047***	(0.007)	0.954	– 0.008	(0.008)	0.992
Age				0.077***	(0.002)	1.080	0.011***	(0.002)	1.011
Age squared				– 0.001***	(0.000)	0.999	– 0.000***	(0.000)	1.000
Middle educated				0.269***	(0.008)	1.309	1.152***	(0.026)	3.165
Higher educated				0.605***	(0.008)	1.831	1.034***	(0.032)	2.812
Education missing				0.446***	(0.008)	1.562	0.809***	(0.018)	2.246
Female				– 0.185***	(0.004)	0.831	– 0.129***	(0.004)	0.879
Months since move				0.000	(0.000)	1.000	– 0.003***	(0.000)	0.997
Correction components included	NO			NO			YES		
Constant	9.871***	(0.025)		7.995***	(0.039)		8.939***	(0.046)	
Residual variance at neighbourhood level	0.009***	(0.001)		0.004***	(0.000)		0.002***	(0.000)	
Residual variance at individual level	0.389***	(0.001)		0.315***	(0.001)		0.230***	(0.001)	
N	84,935			84,935			84,935		

Standard errors in parentheses

*$p < 0.05$

**$p < 0.01$

***$p < 0.001$

Table 5 Neighbourhood effects on income: Rotterdam

	Model 1			Model 2			Model 3		
	B	SE	Exp	B	SE	Exp	B	SE	Exp
Average neighbourhood income (in 10,000 euro)	0.279***	(0.009)	1.322	0.167***	(0.006)	1.182	0.089***	(0.004)	1.093
Non-Western minority				− 0.119***	(0.005)	0.888	1.316***	(0.047)	3.728
Western minority				− 0.115***	(0.006)	0.891	1.299***	(0.032)	3.666
Couple				0.086***	(0.005)	1.090	0.043***	(0.006)	1.044
Couple with children				0.109***	(0.005)	1.115	0.044***	(0.007)	1.045
Other family type				− 0.058***	(0.007)	0.944	− 0.085***	(0.008)	0.919
Age				0.048***	(0.002)	1.049	− 0.003	(0.002)	0.997
Age squared				− 0.000***	(0.000)	1.000	0.000	(0.000)	1.000
Middle educated				0.203***	(0.007)	1.225	− 0.008	(0.017)	0.992
Higher educated				0.527***	(0.007)	1.694	− 0.745***	(0.023)	0.475
Education missing				0.315***	(0.007)	1.370	− 0.093***	(0.017)	0.911
Female				− 0.218***	(0.004)	0.804	− 0.178***	(0.004)	0.837
Months since move				0.001***	(0.000)	1.001	0.002***	(0.000)	1.002
Correction components included	NO			NO			YES		
Constant	9.739***	(0.028)		8.688***	(0.037)		10.046***	(0.038)	
Residual variance at neighbourhood level	0.009***	(0.001)		0.003***	(0.000)		0.001***	(0.000)	
Residual variance at individual level	0.298***	(0.001)		0.237***	(0.001)		0.191***	(0.001)	
N	59,681			59,681			59,681		

Standard errors in parentheses

*$p < 0.05$

**$p < 0.01$

***$p < 0.001$

characteristics not only explain differences between individuals but also between neighbourhoods in individual income.

In Model 3 we added the correction components derived from the previous modelling step in order to control for neighbourhood selection. As we expected, the effect of average neighbourhood income became even smaller than in Model 2. The reduction in effect size between Model 2 and Model 3 is 53% in Utrecht, 52% in Amsterdam and 48% in Rotterdam. These findings indicate that neighbourhood effects on individual income are overestimated to a large degree when the model does not explicitly control for the non-random selection of neighbourhoods by households. The effects of average neighbourhood income remain, however, positive and statistically significant. The exponentiated coefficients from the final model show that in Utrecht a €10,000 difference in average neighbourhood income is related to a 6% difference in individual income in Utrecht, a 8% difference in Amsterdam, and a 9% difference in Rotterdam (on average after 3 years after the move). The neighbourhood effect is significantly weaker in Utrecht compared to Amsterdam and Rotterdam.[6]

Although the effects of the correction components cannot be interpreted unambiguously, and are therefore not reported, Model 3 shows that residential selection plays an important role. A large part of the neighbourhood effect found in Model 1 is the result of residential selection. The inclusion of individual characteristics in Model 2 corrected for a part for this selection; still, the inclusion of the correction components in Model 3 controlled for selection to a larger extent. For all three cities, the residual variances of individual income on the individual level are smaller in Model 3 than in Model 2, showing that the correction components, which were created using interactions between individual and neighbourhood characteristics, explain additional variance in individual income.

Effect Models: Income Change

In addition to the previous analyses of individual income, we provide additional analyses of income *change* between the year of move and 2015 directly (Tables 6, 7, 8). Analysing income change could be seen as a more robust approach to causality, although it is not without problems. In the following models, we keep individual income in 2015 as the dependent variable and include in the models the individual income in the year of move as one of the independent variables. As could be expected, previous income is by far the most important predictor of current income. However, our predictor of interest, the average neighbourhood income, is still significant and follows a similar pattern as in the previous models. The effect of neighbourhood income becomes smaller after including individual characteristics (Model 2) and then even smaller when controlling for selection by including the correction components (Model 3), confirming our predictions. The reduction in

[6] The effect of average neighbourhood income in Model 3 does not significantly differ between Amsterdam and Rotterdam ($Z = \frac{.079-.089}{\sqrt{.004^2+.004^2}} = -1.76$); however, it is significantly weaker in Utrecht compared to Amsterdam ($Z = \frac{.079-.061}{\sqrt{.004^2+.003^2}} = 3.6$), and therefore also Rotterdam.

Table 6 Neighbourhood effects on income change: Utrecht

	Model 1			Model 2			Model 3		
	B	SE	Exp	B	SE	Exp	B	SE	Exp
Average neighbourhood income (in 10,000 euro)	0.013***	(0.002)	1.013	0.011***	(0.002)	1.011	0.010***	(0.002)	1.010
Log income in the year of move	0.963***	(0.003)	2.620	0.935***	(0.003)	2.547	0.923***	(0.004)	2.517
Non-Western minority				− 0.021***	(0.004)	0.979	− 0.017	(0.033)	0.983
Western minority				− 0.018***	(0.004)	0.982	− 0.014	(0.009)	0.986
Couple				0.005	(0.003)	1.005	− 0.002	(0.004)	0.998
Couple with children				− 0.007*	(0.003)	0.993	− 0.040***	(0.005)	0.961
Other family type				0.006	(0.005)	1.006	− 0.008	(0.006)	0.992
Age				− 0.001	(0.001)	0.999	− 0.003*	(0.001)	0.997
Age squared				− 0.000	(0.000)	1.000	0.000	(0.000)	1.000
Middle educated				0.039***	(0.005)	1.040	0.061***	(0.015)	1.063
Higher educated				0.103***	(0.005)	1.108	0.121***	(0.021)	1.129
Education missing				0.051***	(0.005)	1.052	0.062***	(0.015)	1.064
Female				− 0.045***	(0.003)	0.956	− 0.044***	(0.003)	0.957
Months since move				0.002***	(0.000)	1.002	0.001***	(0.000)	1.001
Correction components included	NO			NO			YES		
Constant	0.435***	(0.026)		0.648***	(0.032)		0.876***	(0.043)	
Residual variance at neighbourhood level	0.000***	(0.000)		0.000***	(0.000)		0.000***	(0.000)	
Residual variance at individual level	0.071***	(0.000)		0.068***	(0.000)		0.068***	(0.000)	
N	54,045			54,045			54,045		

Standard errors in parentheses

*$p < 0.05$

**$p < 0.01$

***$p < 0.001$

Table 7 Neighbourhood effects on income change: Amsterdam

	Model 1			Model 2			Model 3		
	B	SE	Exp	B	SE	Exp	B	SE	Exp
Average neighbourhood income (in 10,000 euro)	0.026***	(0.002)	1.026	0.024***	(0.002)	1.024	0.020***	(0.002)	1.020
Log income in the year of move	0.908***	(0.002)	2.479	0.884***	(0.003)	2.421	0.837***	(0.003)	2.309
Non-Western minority				− 0.029***	(0.003)	0.971	0.003	(0.028)	1.003
Western minority				− 0.022***	(0.003)	0.978	− 0.031***	(0.009)	0.969
Couple				0.020***	(0.003)	1.020	0.021***	(0.004)	1.021
Couple with children				0.013***	(0.003)	1.013	− 0.009	(0.005)	0.991
Other family type				0.002	(0.005)	1.002	− 0.013*	(0.006)	0.987
Age				− 0.002	(0.001)	0.998	− 0.006***	(0.001)	0.994
Age squared				− 0.000	(0.000)	1.000	0.000*	(0.000)	1.000
Middle educated				0.059***	(0.005)	1.061	0.170***	(0.020)	1.185
Higher educated				0.147***	(0.005)	1.158	0.255***	(0.024)	1.290
Education missing				0.084***	(0.005)	1.088	0.152***	(0.014)	1.164
Female				− 0.053***	(0.003)	0.948	− 0.054***	(0.003)	0.947
Months since move				0.003***	(0.000)	1.003	0.001***	(0.000)	1.001
Correction components included	NO			NO			YES		
Constant	0.974***	(0.024)		1.144***	(0.030)		1.712***	(0.044)	
Residual variance at neighbourhood level	0.001***	(0.000)		0.000***	(0.000)		0.000***	(0.000)	
Residual variance at individual level	0.136***	(0.000)		0.131***	(0.000)		0.130***	(0.000)	
N	84,935			84,935			84,935		

Standard errors in parentheses

*$p<0.05$

**$p<0.01$

***$p<0.001$

Table 8 Neighbourhood effects on income change: Rotterdam

	Model 1			Model 2			Model 3		
	B	SE	Exp	B	SE	Exp	B	SE	Exp
Average neighbourhood income (in 10,000 euro)	0.034***	(0.003)	1.035	0.025***	(0.002)	1.025	0.022***	(0.002)	1.022
Log income in the year of move	0.903***	(0.003)	2.467	0.868***	(0.003)	2.382	0.838***	(0.004)	2.312
Non-Western minority				− 0.023***	(0.003)	0.977	0.180***	(0.035)	1.197
Western minority				− 0.031***	(0.004)	0.969	0.172***	(0.024)	1.188
Couple				0.009*	(0.004)	1.009	0.001	(0.004)	1.001
Couple with children				− 0.008*	(0.004)	0.992	− 0.029***	(0.005)	0.971
Other family type				− 0.004	(0.005)	0.996	− 0.018**	(0.006)	0.982
Age				− 0.004***	(0.001)	0.996	− 0.007***	(0.001)	0.993
Age squared				0.000**	(0.000)	1.000	0.000***	(0.000)	1.000
Middle educated				0.043***	(0.005)	1.044	0.006	(0.013)	1.006
Higher educated				0.133***	(0.005)	1.142	− 0.011	(0.018)	0.989
Education missing				0.057***	(0.005)	1.059	0.005	(0.012)	1.005
Female				− 0.062***	(0.003)	0.940	− 0.062***	(0.003)	0.940
Months since move				0.002***	(0.000)	1.002	0.002***	(0.000)	1.002
Correction components included	NO			NO			YES		
Constant	0.990***	(0.029)		1.372***	(0.035)		1.767***	(0.046)	
Residual variance at neighbourhood level	0.000***	(0.000)		0.000***	(0.000)		0.000***	(0.000)	
Residual variance at individual level	0.111***	(0.000)		0.106***	(0.000)		0.105***	(0.000)	
N	59,681			59,681			59,681		

Standard errors in parentheses

*$p < 0.05$

**$p < 0.01$

***$p < 0.001$

effect size in Utrecht between Model 1 and 2 is 15%, and between Model 2 and 3 10%. The reduction is 8% and 17% in Amsterdam and 29% and 12% in Rotterdam respectively. Although the reduction in the effect size between Model 2 and 3 is strongest in Amsterdam, the reduction in Utrecht is similar to the reduction for Rotterdam. The exponentiated coefficients from the final model indicate that a one-unit (€10,000) difference in average neighbourhood income is related to a 1% increase in individual income in Utrecht, a 2% increase in Amsterdam, and a 2.2% increase in Rotterdam. The neighbourhood effect is significantly stronger in Amsterdam and Rotterdam compared to Utrecht.[7] We realise that this average income increase is small, yet while interpreting the results one has to remember that our models only include those who have recently moved, and that the time of exposure to their new neighbourhood is relatively short, so large effects cannot be expected.

All other variables follow similar patterns to those in the income models of the previous section, except for the effect of age, which becomes negative (older people's salary is less likely to increase) and months since move, which have a small positive effect on income change in all the models. Also having a partner and children has a negative effect, suggesting that people at this stage of their household careers are less likely to see their income positively change.

Conclusions and Discussion

This study examined how the modelled neighbourhood effect on individual income is altered when controlling for neighbourhood selection, and how these results vary across three Dutch urban regions. Using multilevel models we have estimated neighbourhood effects on income and income change, while controlling for neighbourhood selection correction components. We found that a higher neighbourhood average income was related to a higher individual income, even after controlling for individual characteristics or neighbourhood selection in the form of correction components. The neighbourhood effect becomes even smaller after controlling for selection than after controlling for individual characteristics, which suggests that without taking selection into account, researchers can overestimate neighbourhood effects. The remaining neighbourhood effect is a much smaller, but also a more robust measure of contextual effects on individual income.

The selection model used provide insight into the patterns of neighbourhood selection in three Dutch cities' regional housing markets, with slight local differences and repeated patterns of structured self-sorting, largely in line with the previous studies on the topic. The differences found could provide inspiration to future studies. For example, higher educated individuals tend to select neighbourhoods with a higher percentage of people with non-Western migrant background in Amsterdam and Utrecht, but not Rotterdam; the possible explanations could relate to

[7] The effect of average neighbourhood income on change in individual income (Model 3) is significantly stronger in Utrecht compared to Amsterdam ($Z = \frac{.020-.010}{\sqrt{.002^2+.002^2}} = 3.53$) and not significantly different between Amsterdam and Rotterdam ($Z = \frac{.020-.010}{\sqrt{.002^2+.002^2}} = 3.53$).

different forms of gentrification in these cities, which might lead to different types of neighbourhoods seen as desirable. The selection models also show evidence for difference in preferences based on predictors such as education level, rather than just on earnings (Pinkster & van Kempen, 2002; Jansen, 2012), with higher educated people preferring centrally-located, busy neighbourhoods regardless of household income.

We found clear differences between the three cities in the effects models: the weakest neighbourhood effects can be observed in Utrecht, which also has the highest percentage of native Dutch individuals in our sample, and the strongest in Rotterdam, which has a high percentage of ethnic minorities, just like Amsterdam, but a lower percentage of higher- and middle educated people and a lower average income. This is consistent with the theories of lower-income people being, on average, more vulnerable to negative neighbourhood effects, as they have fewer resources to isolate themselves from the neighbourhood context (Galster et al., 2016); as well as with the studies showing stronger effects in poorer cities (Burdick-Will et al., 2011). However, these results contradict our predictions, based on the assumption that ethnicity and income are to a large extent controlled for in our models. It is possible that stronger effects in Amsterdam and Rotterdam are caused by the influence of general city population, not only the movers included in the model; additionally, difficult to measure characteristics, such as the density of social ties in an average neighbourhood or local policies promoting social cohesion, may be at play.

We also observed differences in the effects of local context-dependent selection mechanisms on modelling neighbourhood effects. In the case of the income models, the reduction in the neighbourhood effect after controlling for selection in Utrecht and Amsterdam was more pronounced than in Rotterdam. This suggests that incorporating neighbourhood selection in neighbourhood effects models might be especially important in higher income cities, in which the competitive nature of the housing market leads to particularly structured selection processes and conscious decisions of the movers. One of the explanations could be that affluent parents support the housing careers of their offspring through social reproduction strategies and the intergenerational transfer of resources (Hochstenbach & Boterman, 2017; Galster & Wessel, 2019), and such strategies are more likely to be employed in richer regions, where a dwelling in the right location is a particularly important investment. However, it is important to note that there might be other explanations for the differences in the attenuation of the effect between the cities, such as differing neighbourhood effect magnitudes because of, for example, local economic conditions which were not captured by the variables in the models. Also, in the income change models it is the reduction in the effect size in Utrecht which was weaker than that in Rotterdam and especially Amsterdam. Future studies could explore these differences further, comparing data from more diverse cities.

Despite the very high quality of the data at our disposal, there are several limitations to our approach. While our models capture neighbourhood selection, some selection might still remain unmeasured, posing a challenge for future research. There might be unobserved variables which could contribute to a better fit of the model, such as more detailed sociocultural predictions, possibly interacting with education (van Gent et al., 2019). Furthermore, people's preferences and therefore neighbourhood selection could be influenced by their prior residential experiences (Bruch & Mare, 2012; Van Ham et al., 2014; Hochstenbach & Boterman, 2017). Including neighbourhood histories could be a next step for future research, as it is beyond the scope of this paper. Another limitation of our approach is that we model neighbourhood effects only for the first couple of years after the move of employed household heads. Especially the short time period might have caused the small sizes of observed effects; this small effect could build over time. We limited our sample to people who moved recently to ensure that we have adequate data from the time of move; but through the use of data reaching further back in time, future studies might be able to compare the neighbourhood selection and neighbourhood effects models for recent "movers" and those who have lived in the neighbourhood for a longer time. Future studies could also use our approach to analyse neighbourhood effects on other outcome measures, such as health or employment, for other groups of people, such as the unemployed and second earners, and for those spending more time at home as they are likely more susceptible to neighbourhood effects (Galster et al., 2016). Also interactions of neighbourhood characteristics with individual characteristics, such as ethnic background, could be tested, following earlier European research (Musterd et al., 2008; Andersson et al., 2014); and the neighbourhood effects on different ethnic and gender groups could be compared (Galster et al., 2010). Using a predictor based on income groups or share of poor neighbours could also prove more accurate than using average income, and has shown interesting results in past studies (Galster et al., 2010; Galster & Turner, 2019). If suitable data becomes available, research could include more information on past neighbourhood histories, modelling the influence of local social interactions since childhood and placing adult episodes in a longer time context. In our study we observed that people are, on average, influenced by the income level of the neighbourhood they move into as adults; however, extending the longitudinal analysis and the model to investigate how not one, but many such episodes shape life outcomes, would lead to a better understanding of spatial inequality both for scientists and policymakers. Our method could be also compared with models using instrumental variables and fixed effects, based on the same dataset (Galster & Hedman, 2013). Also, a transnational study using the same methods with data from different countries could be helpful, since there is evidence from Sweden that when taking selection into account, neighbourhood effect may show to be stronger (Hedman & Galster, 2013).

To conclude, we believe that our research contributes to better understanding of spatial socioeconomic mechanisms. By modelling neighbourhood selection and neighbourhood effects for multiple cities, we shed some light on the locally diverse neighbourhood processes, observing the strongest influence of average neighbourhood income in the relatively poor port city of Rotterdam. Most importantly, we show that the effect of average neighbourhood income on individual income becomes much smaller after controlling for explicitly modelled neighbourhood selection. This result suggests that studies that do not control for neighbourhood selection may overestimate the size of neighbourhood effects, and it could serve as an inspiration for researchers and policymakers to consider residential selection as an integral part of any socio-spatial investigation.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12061-021-09411-5.

Funding The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007–2013)/ERC Grant Agreement n. 615159 (ERC Consolidator Grant DEPRIVEDHOODS, Socio-spatial inequality, deprived neighbourhoods, and neighbourhood effects), as well as from European Union's Horizon 2020 research and innovation programme under Grant Agreement n. 727097 (RELOCAL).

Data availability The data that support the findings of this study are not publicly available due to privacy restrictions of Statistics Netherlands.

Declarations

Conflict of interest The authors have no conflicts of interest to declare. All co-authors have seen and agree with the contents of the manuscript.

Refe+rences

Andersson, R., Musterd, S., & Galster, G. (2014). Neighbourhood ethnic composition and employment effects on immigrant incomes. *Journal of Ethnic and Migration Studies, 40*(5), 710–736.

Boone-Heinonen, J., Gordon-Larsen, P., Guilkey, D. K., Jacobs, D. R., Jr., & Popkin, B. M. (2011). Environment and physical activity dynamics: The role of residential self-selection. *Psychology of Sport and Exercise, 12*(1), 54–60.

Boschman, S., & Van Ham, M. (2015). Neighbourhood selection of non-Western ethnic minorities: Testing the own-group effects hypothesis using a conditional logit model. *Environment and Planning A, 47*(5), 1155–1174.

Boterman, W. R. (2012). Residential mobility of urban middle classes in the field of parenthood. *Environment and Planning A, 44*(10), 2397–2412.

Boterman, W. R., & Bridge, G. (2015). Gender, class and space in the field of parenthood: Comparing middle-class fractions in Amsterdam and London. *Transactions of the Institute of British Geographers, 40*(2), 249–261.

Bruch, E. E., & Mare, R. D. (2012). Methodological issues in the analysis of residential preferences, residential mobility, and neighborhood change. *Sociological Methodology, 42*(1), 103–154.

Burdick-Will, J., Ludwig, J., Raudenbush, S. W., Sampson, R. J., Sanbonmatsu, L., & Sharkey, P. (2011). Converging evidence for neighborhood effects on children's test scores: An experimental, quasi-experimental, and observational comparison. *Whither Opportunity*, 255–276.

Cheah, B. C. (2009). Clustering standard errors or modeling multilevel data. *University of Columbia*, 2–4. https://statmodeling.stat.columbia.edu/.

Endendijk, J. J., Derks, B., & Mesman, J. (2018). Does parenthood change implicit gender-role stereotypes and behaviors? *Journal of Marriage and Family, 80*(1), 61–79.

Galster, G., Andersson, R., & Musterd, S. (2010). Who is affected by neighbourhood income mix? Gender, age, family, employment and income differences. *Urban Studies, 47*(14), 2915–2944.

Galster, G. (2012). The mechanism (s) of neighbourhood effects: Theory, evidence, and policy implications. In *Neighbourhood Effects Research: New Perspectives* (pp. 23–56). Springer.

Galster, G., & Hedman, L. (2013). Measuring neighbourhood effects non-experimentally: How much do alternative methods matter? *Housing Studies, 28*(3), 473–498.

Galster, G., Andersson, R., & Musterd, S. (2016). Neighborhood social mix and adults' income trajectories: Longitudinal evidence from Stockholm. *Geografiska Annaler: Series B Human Geography, 98*(2), 145–170.

Galster, G., & Sharkey, P. (2017). Spatial foundations of inequality: A conceptual model and empirical overview. *RSF: The Russell Sage Foundation Journal of the Social Sciences, 3*(2), 1–33.

Galster, G., & Turner, L. M. (2019). Status aversion, attraction and discrepancy as drivers of neighborhood selection. *City & Community, 18*(3), 937–964.

Galster, G., & Wessel, T. (2019). Reproduction of social inequality through housing: A three-generational study from Norway. *Social Science Research, 78*, 119–136.

Hansen, E. B., & Gottschalk, G. (2006). What makes older people consider moving house and what makes them move? *Housing, Theory and Society, 23*(01), 34–54.

Hedman, L., & Galster, G. (2013). Neighbourhood income sorting and the effects of neighbourhood income mix on income: A holistic empirical exploration. *Urban Studies, 50*(1), 107–127.

Hedman, L., & Van Ham, M. (2012). Understanding neighbourhood effects: Selection bias and residential mobility. In *Neighbourhood Effects Research: New Perspectives* (pp. 79–99). Springer.

Hedman, L., Van Ham, M., & Manley, D. (2011). Neighbourhood choice and neighbourhood reproduction. *Environment and Planning A, 43*(6), 1381–1399.

Hochstenbach, C., & Boterman, W. R. (2017). Intergenerational support shaping residential trajectories: Young people leaving home in a gentrifying city. *Urban Studies, 54*(2), 399–420.

Ioannides, Y. M., & Zabel, J. E. (2008). Interactions, neighborhood selection and housing demand. *Journal of Urban Economics, 63*(1), 229.

Jansen, S. J. (2012). What is the worth of values in guiding residential preferences and choices? *Journal of Housing and the Built Environment, 27*(3), 273–300.

Jencks, C., & Mayer, S. E. (1990). The social consequences of growing up in a poor neighborhood. *Inner-City Poverty in the United States, 111*, 186.

Jokela, M. (2015). Does neighbourhood deprivation cause poor health? Within-individual analysis of movers in a prospective cohort study. *Journal of Epidemiology and Community Health, 69*(9), 899–904.

Jones, K. (1991). Specifying and estimating multi-level models for geographical research. *Transactions of the Institute of British Geographers, 16*(2), 148–159.

Katz, L. F., Kling, J. R., & Liebman, J. B. (2001). Moving to opportunity in Boston: Early results of a randomized mobility experiment. *The Quarterly Journal of Economics, 116*(2), 607–654.

Krysan, M., & Crowder, K. (2017). *Cycle of Segregation: Social Processes and Residential Stratification*: Russell Sage Foundation.

Leventhal, T., & Brooks-Gunn, J. (2003). Moving to opportunity: An experimental study of neighborhood effects on mental health. *American Journal of Public Health, 93*(9), 1576–1582.

McAvay, H. (2018). How Durable Are Ethnoracial Segregation and Spatial Disadvantage? Intergenerational Contextual Mobility in France. *Demography, 55*(4), 1507–1545.

Merton, R. K. (1968). The Matthew effect in science: The reward and communication systems of science are considered. *Science, 159*(3810), 56–63.

Mroz, T. A. (1987). The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions. *Econometrica: Journal of the Econometric Society, 55*(4), 765–799.

Musterd, S., Andersson, R., Galster, G., & Kauppinen, T. M. (2008). Are immigrants' earnings influenced by the characteristics of their neighbours? *Environment and Planning A, 40*(4), 785–805.

Musterd, S., Galster, G., & Andersson, R. (2012). Temporal dimensions and measurement of neighbourhood effects. *Environment and Planning A, 44*(3), 605–627.

Musterd, S., Van Gent, W. P., Das, M., & Latten, J. (2016). Adaptive behaviour in urban space: Residential mobility in response to social distance. *Urban Studies, 53*(2), 227–246.

Oreopoulos, P. (2003). The long-run consequences of living in a poor neighborhood. *The Quarterly Journal of Economics, 118*(4), 1533–1575.

Paternoster, R., Brame, R., Mazerolle, P., & Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology, 36*(4), 859–866.

Permentier, M., Van Ham, M., & Bolt, G. (2008). Same neighbourhood… different views? A confrontation of internal and external neighbourhood reputations. *Housing Studies, 23*(6), 833–855.

Phinney, R. (2013). Exploring residential mobility among low-income families. *Social Service Review, 87*(4), 780–815.

Pinkster, F. M. (2007). Localised social networks, socialisation and social mobility in a low-income neighbourhood in the Netherlands. *Urban Studies, 44*(13), 2587–2603.
Pinkster, F. M. (2014). Neighbourhood effects as indirect effects: Evidence from a Dutch case study on the significance of neighbourhood for employment trajectories. *International Journal of Urban and Regional Research, 38*(6), 2042–2059.
Pinkster, F. M., & van Kempen, R. (2002). *Leefstijlen en woonmilieuvoorkeuren: Urban and Regional Research Centre Utrecht*. Universiteit Utrecht.
Rossi, P. H. (1955). *Why families move: A study in the social psychology of urban residential mobility*: Free Press.
Sampson, R. J. (2008). Moving to inequality: Neighborhood effects and experiments meet social structure. *American Journal of Sociology, 114*(1), 189–231.
Sampson, R. J. (2012). *Great American city: Chicago and the enduring neighborhood effect*: University of Chicago Press.
Sampson, R. J., & Laub, J. H. (1997). A life-course theory of cumulative disadvantage and the stability of delinquency. *Developmental Theories of Crime and Delinquency, 7*, 133–161.
Sampson, R. J., & Sharkey, P. (2008). Neighborhood selection and the social reproduction of concentrated racial inequality. *Demography, 45*(1), 1–29.
Sanbonmatsu, L., Kling, J. R., Duncan, G. J., & Brooks-Gunn, J. (2006). Neighborhoods and academic achievement results from the Moving to Opportunity experiment. *Journal of Human Resources, 41*(4), 649–691.
Sari, F. (2012). Analysis of neighbourhood effects and work behaviour: Evidence from Paris. *Housing Studies, 27*(1), 45–76.
Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology, 1*(2), 143–186.
Small, M. L., & Feldman, J. (2012). Ethnographic evidence, heterogeneity, and neighbourhood effects after moving to opportunity. In *Neighbourhood Effects Research: New Perspectives* (pp. 57–77): Springer.
Stead, D., & Meijers, E. J. (2015). Urban planning and transport infrastructure provision in the Randstad, Netherlands: A Global City Cluster. *Joint ITF-TPRI Roundtable on Integrated Transport Development Experiences Of Global City Clusters, Beijing, China, 2–3 July 2015*.
Toft, M., & Ljunggren, J. (2016). Geographies of class advantage: The influence of adolescent neighbourhoods in Oslo. *Urban Studies, 53*(14), 2939–2955.
van Gent, W., Das, M., & Musterd, S. (2019). Sociocultural, economic and ethnic homogeneity in residential mobility and spatial sorting among couples. *Environment and Planning A: Economy and Space, 51*(4), 891–912.
Van Ham, M., Boschman, S., & Vogel, M. (2018). Incorporating Neighborhood Choice in a Model of Neighborhood Effects on Income. *Demography, 55*(3), 1069–1090. https://doi.org/10.1007/s13524-018-0672-9
Van Ham, M., Hedman, L., Manley, D., Coulter, R., & Östh, J. (2014). Intergenerational transmission of neighbourhood poverty: an analysis of neighbourhood histories of individuals. *Transactions of the Institute of British Geographers, 39*(3), 402–417. https://doi.org/10.1111/tran.12040.doi:10.1111/tran.12040.
Van Ham, M., Hooimeijer, P., & Mulder, C. H. (2001). Urban form and job access: Disparate realities in the Randstad. *Tijdschrift Voor Economische En Sociale Geografie, 92*(2), 231–246.
Wilson, W. J. (1987). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*: University of Chicago Press.
Winship, C., & Mare, R. D. (1992). Models for sample selection bias. *Annual Review of Sociology, 18*(1), 327–350.
Zick, C. D., Hanson, H., Fan, J. X., Smith, K. R., Kowaleski-Jones, L., Brown, B. B., & Yamada, I. (2013). Re-visiting the relationship between neighbourhood environment and BMI: An instrumental variables approach to correcting for residential selection bias. *International Journal of Behavioral Nutrition and Physical Activity, 10*(1), 27.