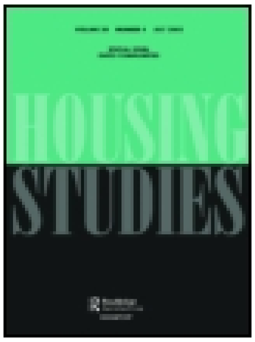
Article:

Bakens, J. and Pryce, G.B. (2019) Homophily horizons and ethnic mover flows among homeowners in Scotland. Housing Studies, 34 (6). pp. 925-945. ISSN: 0267-3037

https://doi.org/10.1080/02673037.2018.1504007



Routledge
Taylor & Francis Group

Housing Studies

ISSN: 0267-3037 (Print) 1466-1810 (Online) Journal homepage: http://www.tandfonline.com/loi/chos20

Homophily horizons and ethnic mover flows among homeowners in Scotland

Jessie Bakens & Gwilym Pryce

HOUSING STUDIES
https://doi.org/10.1080/02673037.2018.1504007

Routledge
Taylor & Francis Group

RESEARCH-ARTICLE

OPEN ACCESS Check for updates

Homophily horizons and ethnic mover flows among homeowners in Scotland

Jessie Bakens[a] and Gwilym Pryce[b]

[a]Research Centre for Education and the Labour Market, Maastricht University, Maastricht, The Netherlands; [b]Sheffield Methods Institute, University of Sheffield, Sheffield, United Kingdom

ABSTRACT

This article analyses mover flows in Glasgow and the role of ethnic homophily, the tendency for movers to be drawn to areas with similar ethnicities to their own. We look at how homophily affects the spatial relocation patterns of homeowners in Glasgow from Scottish, Indian, Pakistani and Chinese descent, and focus on the extent to which homophily extends beyond the immediate locality to surrounding neighbourhoods. Our interest is in estimating the "homophily horizon" – how far the gaze of homophily reaches in mover location decisions. Using a simple Schelling-type theoretical model, we argue that homophily horizons are potentially important in shaping the long-term social structure of cities as they may profoundly affect how potent the overall sorting tendencies of the housing market are in driving segregation. In principle, the more distant the homophily horizon, the more quickly the housing market will tend towards segregation, other things being equal. We adopt Folch and Rey's use of the local centralization index to capture the influence of surrounding neighbourhoods in shaping mover flows and neighbourhood dynamics. Our estimation combines ethnic mover flows derived from surname analysis of house buyers from the house transactions recorded in Registers of Scotland data. Our results show that the presence of the own ethnic group in the local surroundings is important for explaining mover flows, and that homophily is a local phenomenon.

ARTICLE HISTORY
Received 6 October 2016
Accepted 12 July 2018

KEYWORDS
Neighbourhoods; home ownership; ethnic location decisions; migration

1. Introduction

The role of an individual's ethnic background in residential location outcomes has long been the subject of research in many academic fields across many countries. One of the approaches in this very broad field of research is the analysis of preferences of individuals to reside in neighbourhoods with their own ethnic group, and, consequently, the preference to reside in neighbourhoods with ethnic groups different from their own ethnic group. The ethnic composition of neighbourhoods and the distribution of ethnic groups among neighbourhoods in an area (i.e. city), then drives

CONTACT Jessie Bakens j.bakens@maastrichtuniversity.nl

location dynamics which in turn determines the degree of segregation in the housing system as a whole (Becker & Murphy, 2000; Schelling, 1971).

Segregation patterns in neighbourhoods reflect not only the difference in socioeconomic status of ethnic minorities compared to the majority group, but also differences among ethnic minorities themselves. The economic position of many immigrants and ethnic minority groups makes them dependent on low cost (social) housing, which limits location choice. However, in many studies, these socioeconomic differences between natives and ethnic minorities cannot fully explain the concentration and clustering of ethnic minority groups. Preferences for the own ethnic group in residential location choice is found to be part of the explanation (Bayer *et al.*, 2014; Krysan & Farley, 2002), although the underlying reason for these preferences can be multi-faceted.

Social network-based research points toward the benefits of being close to the own ethnic group in terms of labour market opportunities and other externalities from social networks (see, for example, Borjas, 1995; Portes & Sensenbrenner, 1993). The attractiveness of these ethnic enclaves can have an economic dimension because social networks can improve employment chances (Borjas, 1995). If language proficiency is an obstacle for immigrants, living in ethnic enclaves close to the own immigrant group may be beneficial for job market purposes (Dustmann & Fabbri, 2003). In addition, the amenities in neighbourhoods mirror local demand, so churches, shops, and other amenities that are ethnic-specific are more likely to be present in a neighbourhood with many people of that ethnicity (Waldfogel, 2008).

How minority groups respond to, and are received by, the majority group may also be important. Krysan & Farley (2002), for example, describe the fear of hostility experienced by individuals from ethnic minority groups living in neighbourhoods with very low shares of their own group. There is solid evidence to suggest that such fears are well founded. Recent research by Nandi & Luthra (2016) using UK data finds that for many ethnic minorities, more than 10% of men had been physically or verbally attacked in the last 12 months in public transport, shops, street, other places due to their ethnicity, nationality, or religion. Moreover, the 'likelihood of experiencing ethnic and racial harassment is lower for those living with a higher proportion of their own ethnic group members' (p.1). The tendency for particular minority groups to be drawn to areas with high concentrations of their own ethnicity may therefore reflect the lack of acceptance of that ethnic group in wider society.

Residential clustering may also reflect fundamental beliefs in particular religions and cultures to live separately from wider society. However, these beliefs themselves may have arisen as a response to historical persecution arising from attempts by dominant groups to assimilate or eradicate them. Various strands of social theory have explained the veracity of isolation and segregation as a strategy for preserving group identity. For example, Rytina & Morgan's (1982, p. 88) theoretical synthesis of 'Blau's axiomatic theory of social structure with the quantitative approach of social network analysis,' highlights the importance of isolation as a way of minority groups to maintain their coherence. Group identity and coherence are, in turn, important factors in the formation and preservation of community and solidarity (Giuffre, 2013).

There are good theoretical reasons, therefore, to expect residential choice to be driven by the desire to live near likewise neighbours (see, for example, Bayer *et al.*, 2004, 2014; Ioannides & Zabel, 2008), not only in terms of socio-economic position but also ethnicity and demography (Hedman *et al*., 2011). As the ethnic minority groups in many populations increases, this may lead to more diverse neighbourhoods on the aggregate level, but it can also lead to more segregation. Card *et al*. (2008) show that, with an increasing share of non-whites in US cities, there is 'white flight' and neighbourhoods 'tip'. The share of non-whites that start the white flight process depend on tolerance levels in different cities. Bayer *et al*. (2014), for example, show in the US that if more black households increase their socio-economic position, it becomes possible and preferable for such households to form middle income black neighbourhoods.

In the social network literature, the propensity for location decisions to be driven by preference living close to one's own ethnicity is described as homophily (McPherson *et al*., 2001), analogous to the tendency of birds of a feather to flock together. McPherson *et al*. (2001) describe homophily as '… the principle that a contact between similar people occurs at a higher rate than among dissimilar people.' Homophily is an organizing principle of networks in which individuals with similar characteristics are connected or closer within a network than individuals that are less similar. Schelling (1971) showed that segregated communities (along ethnic lines) can be a stable outcome of location decisions if individuals follow their preferences for the local population composition, all other things equal. In the end, describing segregation patterns and the underlying drivers of these patterns is an empirical issue that differs per city size and the ethnic composition.

A surprisingly under-researched aspect in these large and overlapping literatures is how far the gaze of homophily reaches in mover location decisions. We label this distance the 'homophily horizon,' defined as the extent to which the homophily effect extends beyond the immediate neighbourhood to encompass surrounding areas. The extent of the homophily horizon is potentially important as it will likely effect the long-term social-spatial structure of cities.

To illustrate the theoretical implications for urban social-spatial structures of varying the 'homophily horizon' in a Schelling-type (1971) conceptual framework, we construct a simplified 'minority-mover' model[1] and vary the depth of the cell perimeter considered by residents when deciding whether to move or stay. The simple theoretical model is represented graphically in Figure 1. As is traditional in this genre of models, we assume a grid square of cells (which represent dwellings or clusters of dwellings) inhabited by two types of household, denoted by black and white shading. In round 1, the two household types are distributed randomly. In subsequent rounds, households move out if they find themselves in a neighbourhood where they are in the minority.

We can see from Figure 1 what the impact would be of varying the homophily horizon in a minority-mover model. Here the homophily horizon is defined in terms of the radius of the area (depth of cells around the location in question) considered by residents when deciding whether to move or stay. Panel (b) shows the distribution of black and white households after 40 rounds of relocation choices assuming a neighbourhood radius of 1—i.e. the horizon only has a depth of one cell around the location in question. We can see from panel (c) how the same starting allocation

Figure 1. Impact of homophily horizon on Schelling model outcomes (a) Round 1 random cell assignment (b) Round 40 with neighbourhood radius =1 cell (c) Round 40 with neighbourhood radius =2 cells. 50 × 50 grid square. Source: Own calculations based on Schelling (1971).

leads to a much more segregated urban landscape if we extend the homophily horizon to two cells. This very simple hypothetical simulation illustrates how in principle, at least, the pattern and degree of urban segregation could be sensitive to the homophily horizon. The reality, of course, is likely to be more complex because of the interplay of multiple ethnic groups, each with potentially different degrees of homophily and different homophily horizons, combined with varying spatial structures of social mix within a particular homophily radius. Nevertheless, the basic insight of our hypothetical model is a powerful one. If homophily horizons are relatively near in their spatial reach, the dynamics of the overall housing system will be more 'tolerant' in allowing pockets of different ethnic groups to persist in fairly close proximity. In contrast, ethnic mix is much less stable in a world where homophily horizons extend to surrounding neighbourhoods.

In this article we attempt to explore the idea of homophily horizon while taking into account some of these complexities by analysing relocation patterns of homeowners in Glasgow. We use a partial gravity model of mover flows based on Bakens *et al.* (2018) and focus on inflows into neighbourhoods.[2] We aim to describe (i) at what spatial distance (i.e. horizon) homophily is most potent, (ii) whether this differs between ethnic[3] groups (i.e. ethnic asymmetry), and (iii) whether these spatial segregation structures are important for neighbourhood perceptions and dynamics (i.e. using these indices to explain inflows). The focus on homeowners has the advantage that the observed location choice is taken in a relatively free market with budget and housing availability constraints. This is not the case when looking at, for example, social housing, where role of choice at the level of the household is more limited. Note, however, that we do not analyse individual location choices, nor are we seeking to develop an equilibrium sorting model (see Kuminoff *et al.*, 2013). Rather, we are interested in the main drivers of aggregated inflows of house owners by ethnic background into neighbourhoods. There is not much research focused on ethnic location dynamics in Scotland and Glasgow (McGarrigle (2010) is an exception). The perception of what homophily at a particular spatial level is, might be simple or more spatially complex. Research on segregation has introduced many measurements that each highlight a different aspect of established spatial segregation patterns (Massey & Denton, 1988). It might simply mean that there are many people who are ethnically alike, i.e. the share of the own ethnic group is relatively high in that location. Or it might mean that there is a degree of spatial ordering of ethnicity with respect to particular ethnic hubs. For example, a large part of an ethnic group is located in or around that neighbourhood, resulting in that neighbourhood being more or less central to the overall distribution of the own ethnic group across neighbourhoods. Moreover, the spatial ordering of ethnicity relative to that ethnic hub might be a critical factor in shaping ethnic mover flows into particular neighbourhoods.

Our underlying theoretical rationale translates these spatial notions into homophily processes, interpreting them in terms of hypothetical network connections and network position. We assume that spatial exposure to the own ethnic group will reflect the number of potential connections in a neighbourhood. Regarding network position, we are interested in the ethnic centralization of a neighbourhood in relation to the total neighbourhood structure of a city which serves as an indicator of the position of that neighbourhood in the city-wide ethnic network. That is, we seek to measure the extent to which a neighbourhood is at the centre or periphery of the spatial ordering of neighbourhoods with respect to ethnic enclaves. We attempt to capture these notions of network connection and network position using ethnic share in a neighbourhood and local centralization measures, respectively. With respect to the latter, Folch & Rey (2015) propose a measure in which the centralization of a random spatial unit can be calculated to indicate how minority and majority populations are spatially ordered with respect to that spatial unit. By introducing a non-parametric inference of the local centralization measure, Folch (2012) and Folch & Rey (2015) provide a way of identifying a plausible selection of indices based on statistical significance.

By mapping and calculating the local centralization index including the inference for different ethnic groups, we can shed light on the patterns of concentration and centralization of ethnic minority groups and the spatial persistence of these patterns over time. The larger the relative mover inflows of an ethnic group into a neighbourhood, the larger the pull of homophily.[4] The effect of homophily, both the concentration and centralization is likely to be nonlinear: Schelling (1971), Krysan & Farley (2002) and Card *et al.* (2008) all show that there are tipping points at which neighbourhoods become very (un)attractive for specific groups once the population has a certain composition. In our analysis, the nonlinearity can be included in the analysis, as the effect of concentration and centralization on the size of the mover flow is not restricted to be linear. The results could show an increase of ethnic enclaves where there is a strong concentration of specific ethnic groups and homophily plays a large role. This result will be more pronounced if centralization plays an important role as well and the homophily horizon is larger. The opposite result may be described as spatial assimilation (see, for example, Alba *et al.*, 1999), where homophily plays a small role in explaining mover flows and neighbourhoods become more ethnically diverse. In the latter case, socio-economic status might be more important: ethnic minority groups may choose more ‘White’ neighbourhoods when they increase their economic status (Bayer *et al.*, 2014; Massey & Denton, 1985).

Our results for Glasgow suggest that homophily is important for explaining mover flows, especially the presence of the own ethnic group in the local surroundings. We find that the homophily horizon is relatively short, indicating that it is the ethnic mix of the immediate locality, rather than the spatial ordering of ethnicity in the surrounding neighbourhoods, that dominates homophily behaviour. In line with Bakens *et al.* (2018) for Amsterdam and the Hague in the Netherlands, the role of the own ethnic group in the neighbourhood seems important for the size of ethnic mover flows.

In the next section we set out the identification strategy used in our analysis. We use the Data section to explain how we infer the name matching of the house buyers with observed ethnicity from the Scottish census data. The ethnicity of house buyers is identified by (sur)name analysis (Mateos, 2014), which can overcome some of the drawbacks inherent in using other ways of identifying ethnicity, such as country of birth, for example. There is a rich literature on using Onomastics to identify ethnicity. In the Results section, we map the centralization indices and give the results of our analysis of mover flows. The final section concludes and reflects on future directions of research.

2. Identification strategy

We use a combination of nonspatial and spatial segregation measures to describe the role of ethnic homophily in location decisions. We measure local exposure of ethnic group x to other groups by the simple nonspatial share of group x in the total population of a neighbourhood[5]:

$$P_{\mathrm{xi}} = \frac{x_{\mathrm{i}}}{t_{\mathrm{i}}}, \tag{1}$$

with x_{i} denoting the size of the population of ethnic group x in neighbourhood i, and t_{i} the total population of neighbourhood i. A true exposure measure (Massey & Denton, 1988) is generally calculated for a whole city, however our identification strategy requires variation over the units of measurement. The share of an ethnic group in a neighbourhood has the same underlying idea of the true exposure measure, which is whether people from an ethnic group are more likely to meet people from their own group (a high share) or from other groups (a low share).

Where the share focuses on the composition of the neighbourhood, the centralization index focuses on the overall pattern of neighbourhoods in a spatial structure, like a city. Following Folch & Rey (2015), local centralization indices can be calculated for any given area as a centre (as opposed to the city centre or central business district as the main centre). The index then shows the distribution of a specific ethnic population around the central neighbourhood. Folch & Rey (2015) define the local centralization index, C, for neighbourhood i and spatial structure k as:

$$C_{\mathrm{ik}} = \sum_{j=2}^{J} \hat{X}_{j-1}\hat{Y}_j - \sum_{j=2}^{J} \hat{X}_j\hat{Y}_{j-1}, \tag{2}$$

with vectors $\hat{X}$ and $\hat{Y}$ defined as:

$$\begin{aligned}\hat{X} &= \frac{x_1}{X_k}, \frac{x_1 + x_2}{X_k}, \frac{x_1 + x_2 + x_3}{X_k}, \ldots, \frac{x_1 + \cdots + x_{J-1}}{X_k}, 1 \\ \hat{Y} &= \frac{y_1}{Y_k}, \frac{y_1 + y_2}{Y_k}, \frac{y_1 + y_2 + y_3}{Y_k}, \ldots, \frac{y_1 + \cdots + y_{J-1}}{Y_k}, 1.\end{aligned} \tag{3}$$

In Equations (2) and (3), x_j is the number of people from ethnic group X in observation j. The observations are ordered by distance from the centre i. The spatial structure k consists of J spatial units surrounding the centre i. The same definitions hold for the majority population Y; y_j is the number of people from the majority ethnic group Y in observation j. The value of the local centralization index lies between -1 and $+1$, with a positive value indicating centralization of group X around the centre, and a negative value indicating centralization of group Y around the centre. In the empirical analysis, X in Equation (2) is one of the ethnic groups, while Y is defined as $1 - X$, i.e. all the other ethnic groups. We use the ethnic population composition from the 2001 Census to calculate the local centralization index (LCI).

Because this index uses cumulative population shares by ordering observations based on distance to the centre, the index actually encompasses different spatial relations: (i) the location of ethnic group X relative to the chosen centre, (ii) the location of the other group Y relative to the chosen centre, and (iii) the location of group X relative to group Y (Folch & Rey, 2015). The spatial structure k needs to be selected carefully as it drives the overall results. We choose k as the number of neighbourhoods of which the centroids are located within a significant distance from i. The significant distance can be determined by following the inference technique from Folch & Rey (2015) which focusses on whether the distribution of ethnic groups over neighbourhoods is geographically statistically significant, i.e. not random. We use a different way of measuring the inference of the distance k by focussing on the significance of the LCI in explaining the size of mover flows as is explained below. The model-fit of a regression analysis with different radius distances tells us which distance explains mover flows best.

The inflow f of house owners of a specific ethnic group x into a neighbourhood i for a given 5-year period (2001–2005 following the census year 2001) is given by:

$$f_{xi} = F(\beta_0 + \beta_1 P_{xi} + \beta_2 C_{xi} + z'_i \beta_3 + \delta_i), \tag{4}$$

with P_{xi} the share of group x in neighbourhood i, C_{xi} the centralization index for neighbourhood i for ethnic group x, z'_i a vector of neighbourhood characteristics, δ_i unobserved neighbourhood characteristics at the intermediate datazone level, and F the functional form of the regression. The neighbourhood control variables include the median house price and dummies at the intermediate datazone level.

Equation (4) is estimated by a count data model with zeros if there is no inflow of a specific ethnic group into a neighbourhood (see Santos Silva & Tenreyro (2006); Bakens *et al.*, 2018). The probability of observing a number of specific counts is given by a Poisson probability function (Long, 1997; Winkelmann, 2008). The data of most socioeconomic events is not Poisson distributed and the assumption of equidispersion (the events occur independent) is rejected in favour of overdispersion. We therefore estimate a Negative Binomial model.

3. Data

This research is based on two data sources, the Register of Scotland (RoS) house transaction data and the Scottish census data over 2001 and 2011. The house transaction data contains the transactions of houses in Scotland between 1990 and 2010. Details include the exact location of the dwelling, the date and price of the transaction, the names of the sellers and buyers, and information on the previous residential location of the buyers at an aggregate spatial level.[6] We select the observations for which the full postcode details are available, and which have sold at a price between £10,000 and £1,000,000 in real 2010 prices.

The analysis is further focused on the urbanized area of the travel to work area (TTWA) of Glasgow. Ethnic minorities in Scotland are concentrated in the few large cities. The population of Scotland consists of about 84% ethnic white Scotts, and about 4% non-white ethnic minorities according to the 2011 census. Glasgow has the highest share of ethnic minorities in Scotland, and the 2011 census shows that about 17% of the population of the area under consideration in this analysis is of a non-white Scottish background. The Chinese, Indian and Pakistani population constitutes 5.4% of the total population in Glasgow. Focusing on the whole country is therefore not feasible in terms of the variability in the data needed to identify the associations brought forward in this article.

To derive the ethnicity of the house buyers in the RoS dataset, the ONOMAP[7] name-classification system is used (Mateos *et al.*, 2007; Mateos, 2014). ONOMAP identifies the most likely cultural-ethnic-linguistic (CEL) origin of a name, which can be based on surname, forename, or the combination of surname and forename. ONOMAP covers over 500,000 forenames and 1 million surnames from 28 countries. The underlying database of this program is constructed by analysing names of known ethnic, cultural and linguistic groups and the inference of the names within these groups. Names are allocated to one of 185 CEL categories with a precision level depending on whether only surnames are used or whether a sur- and forename are of the same CEL-group or belong to different CEL-groups. Throughout the paper, someone's ethnicity is thus derived from the name-analysis.

We only use observations of private buyers with a CEL-identified category. Our name matching is based on forename and surname when available, and otherwise on surname only. Observations that have more than two buyers of one dwelling are not included, and when there are two buyers identified, only the name of the first mentioned buyer is used. This is mostly the male name in a couple. As the RoS dataset was not designed for name matching, it is not always clear which is the forename and which is the surname. We matched the name twice, with each of the two names as a forename (surname). The highest CEL-score is then taken to identify which is the forename and which is the surname that identifies the CEL-category.[8]

We link the RoS data to the census data to obtain the population composition at the datazone level for the census years 2001 and 2011. As the census only contains a

Table 1. House buyers by ethnic name group 2001–2005.

	2001–2005	
	Frequency	Percentage
Scottish	46,934	36.83
Indian	2,255	1.77
Pakistani	4,873	3.82
Chinese	901	0.71
Other	72,465	56.87
Total	127,428	100.00

Source: Own calculations based on RoS house transaction data.

few distinctive ethnic groups, our analysis is focused on the Scottish, Indian, Pakistani and Chinese ethnic groups in Glasgow. Furthermore, we use the datazone classification from the 2001 census. The TTWA of Glasgow in our analysis consists of 1,158 datazones.[9] The regression analysis then consists of (1158×4) 4632 observations, i.e. 1158 datazones for each ethnic group in our analysis.

From the RoS dataset, 54,952 observations between 2001 and 2005 are name-matched to be house owners in Glasgow with a name originated from Scotland, India, Pakistan, or China. Table 1 shows the number of observations in the analysis between 2001 and 2005. The number of house buyers is aggregated over 5 years (2001–2005) to obtain substantial inflows. This table shows that the largest ethnic minorities in Scotland only constitute a small part of the house buyers in the greater area of Glasgow. The census differentiates between ethnic groups from different parts of the UK, e.g. between Scottish, Other British and Irish respondents. Therefore, the ethnic Scottish population might be perceived as low. Figure 2 maps the datazone shares of each ethnic group from the 2001 census. The scales of the ethnic groups differ, and it is clear that the overall share of each ethnic minority group is small in most datazones. The Indian and Pakistani population seem to be more clustered in the South of the city centre, while the Chinese population is more clustered North of the city centre.

Figure 3 maps the inflow of house owners per ethnic group between 2001 and 2005 as share of the total inflow of house owners in that datazone. The figures show that inflows of house owners with a Pakistani or Indian background are a bit higher in the South of Glasgow than elsewhere, and that the inflows increase in the suburban areas in the South. Inflows of Chinese house owners are higher in some specific datazones in the North part of Glasgow. Although we observe a high number of inflow of people from a specific ethnic group in areas, this can both mean that there is indeed an inflow of new people, but also moving of people within the same datazone. Because we do not know the datazone that people move out of, we cannot distinguish between the two.

The number of house owners from an ethnic group, the inflow, is our dependent variable. Tables 2 and 3 give the descriptives of the dependent and independent variables, respectively. We include all datazones in the urban areas of the TTWA of Glasgow, even if the inflow of a specific ethnic group is zero. Note that the number of houses sold over the period is used as an exposure variable in the Negative Binomial model, and derived from the Scottish Neighbourhood Statistics.

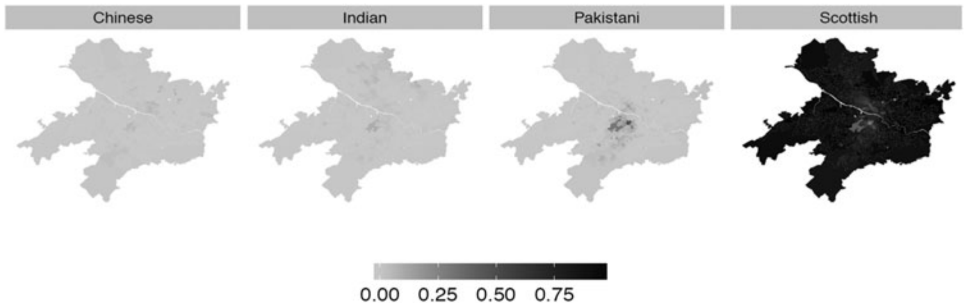


Figure 2. Shares per ethnic group, 2001, Datazones. Source: Own calculations based on Scottish Census 2001.

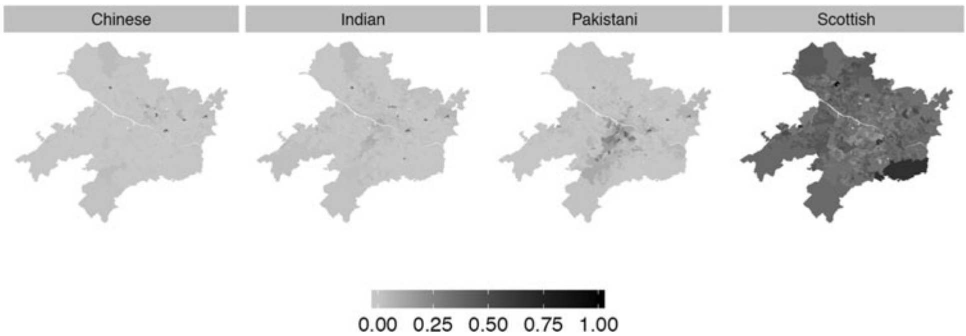


Figure 3. Share of inflow per ethnic group, 2001–2005, Datazones. Source: Own calculations based on Scottish census 2001.

Table 2. Flow size by ethnic group 2001–2005.

	2001–2005			
	Mean	Standard deviation	Minimum	Maximum
Scottish	40.53	28.76	0	216
Indian	1.95	3.34	0	44
Pakistani	4.21	9.12	0	89
Chinese	0.78	1.82	0	23

Source: Own calculations based on RoS transaction data.

4. Results

Folch & Rey (2015) show that the choice of the radius is critical for the results of the LCI and propose an inference of both the radius as well as the calculated centralization indices. The inference of the radius as proposed by Folch & Rey (2015) is based on assumptions about the random distribution of ethnic groups across neighbourhoods in cities. To determine the statistically significant radius the LCI is calculated at increments of the radius by 250 m between 250 and 10,000 m based on 500 permutations for each radius. With each permutation, the LCI is calculated by randomly allocating the Glasgow ethnic neighbourhood composition to the Glasgow neighbourhoods. If the real centralization index is statistically significantly different at a 5% level from the distribution of LCIs based on 500 permutations, the index is considered to be not random (see Folch & Rey (2015) for more detailed information).

Table 3. Independent variables.

	Mean	Standard deviation	Minimum	Maximum
Share 2001				
Scottish	0.89	0.10	0.24	0.98
Indian	0.01	0.01	0	0.11
Pakistani	0.02	0.05	0	0.59
Chinese	0.01	0.01	0	0.16
Centralization Index (share 2001)[a]				
Scottish	0.03	0.17	−0.34	0.39
Indian	−0.06	0.19	−0.53	0.29
Pakistani	−0.10	0.32	−0.66	0.60
Chinese	−0.04	0.18	−0.39	0.36
Median house price (x £10,000)	9.41	5.49	0.50	41.84
No. of sales[b]	88.00	70.81	1	556

Source: Own calculations based on Scottish census data 2001, and RoS transaction data.
[a]The centralization index is calculated for a distance band of 8 km around the centroid of each datazone.
[b]The total number of houses sold is used as the exposure variable in the regression.

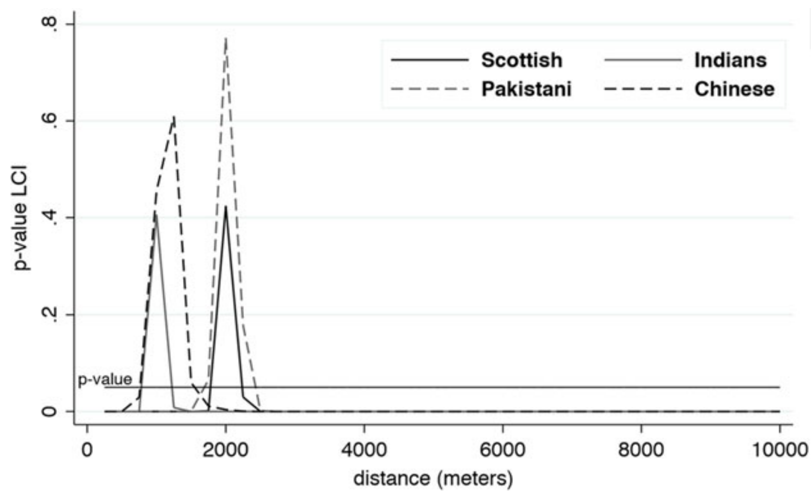


Figure 4. Inference according to Folch & Rey (2015). Source: own calculations based on Scottish census 2001.

As can be seen in Figure 4, the distance at which the LCIs are statistically significant at the 5%-level and not too volatile, is just below 2 km for Chinese and Indian house owners, and just above 2 km for Pakistani and Scottish house owners.

A different way of interpreting the inference of the LCI, is by comparing the model-fit of estimating Equation (4) at different distances for the radius. In the same way as the inference using the permutations, we estimate Equation (4) with the LCI at incremental steps of 250 m between 250 and 10,000 m and compare the model-fit (log-pseudolikelihood) of each estimation. Figure 5 shows that incrementing the distance for all ethnic groups, the best model-fit is at 8 km. If we subsequently look at the model-fit for each ethnic group, keeping the LCI for the other ethnic groups constant at 8 km, we find that except for Scottish house owners, the best fit of the model, i.e. most statistically significant LCI for explaining the size of mover flows, is at 8 km. For Scottish house owners, the model fit is more or less the same at all distances, as the LCI for this group is insignificant at all distances. Scottish house

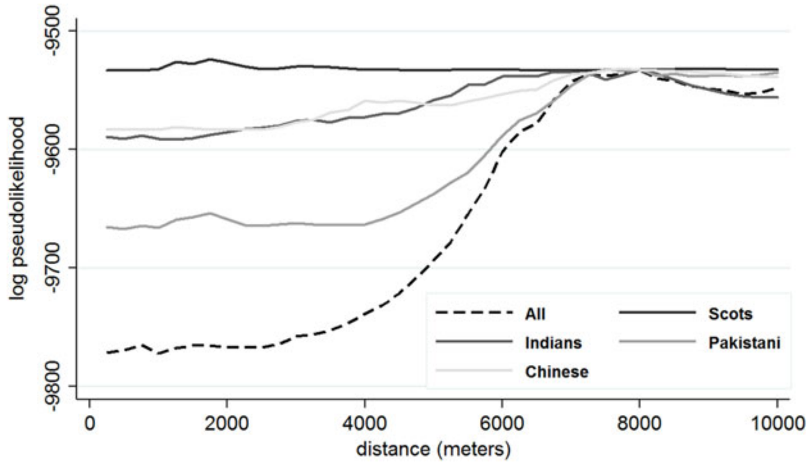


Figure 5. Best fit of regression analysis. Source: Own calculations based on Scottish census 2001.

owners are centralized and numerically dominant in most neighbourhoods in Glasgow, thus the spatial aspect of homophily should matter less in determining mover flows of house owners with a Scottish descent.

Because we are interested in explaining the role of homophily for mover patterns, we will use the results of statistical inference described above to fix *k* at the 8 km radius for the LCI. We use Folch & Rey's (2015) approach to calculate the inference of the LCI at 8 km by permuting the neighbourhood patterns 5,000 times. Figures 6 and 7 map the LCI that are statistically significant at the 0.05 level for 2001 and 2011, respectively. Pockets of statistically significant centralization indices signify areas that are centres of ethnic clustering. We clearly see a concentration of such centres of ethnic minority groups in the inner city circle of Glasgow and to the North and South of the city centre, while Scottish house owners tend to have spatial ordering with respect to a ring of centres around the outskirts of the city. The cluster of statistically significantly segregated Scottish neighbourhoods has declined between 2001 and 2011. There are some small changes in the neighbourhoods with a statistically significant concentration of ethnic groups, especially towards the suburbs. Towards the North- and South-West, centralization of Scottish people has become statistically significant, while centres of spatial ordering of the Indian, Chinese and Pakistani population in Glasgow are gradually more concentrated towards the South-East.

The clustering of ethnic centres of the smaller ethnic groups in Figures 6 and 7 shows a moderate suburbanization pattern. The suburbanisation trend of immigrant groups is also found in the US (see, for example, Massey and Tannen, 2017). Both in Glasgow, and in the research described by Massey and Tannen (2017) for the US, the segregation remains high. The suburbanisation trend is in line with McGarrigle (2010), who gives a detailed description of the residential location patterns of the South Asian population in Glasgow. She also finds that the South Asian population tends to get more dispersed around some of the 'centres' of clustering, which we confirm with these maps. McGarrigle (2010) also finds that the Indian minorities tend to spatially cluster less than the Pakistani minorities. McGarrigle (2010) also uses qualitative data to describe underlying motives for location choices, and points

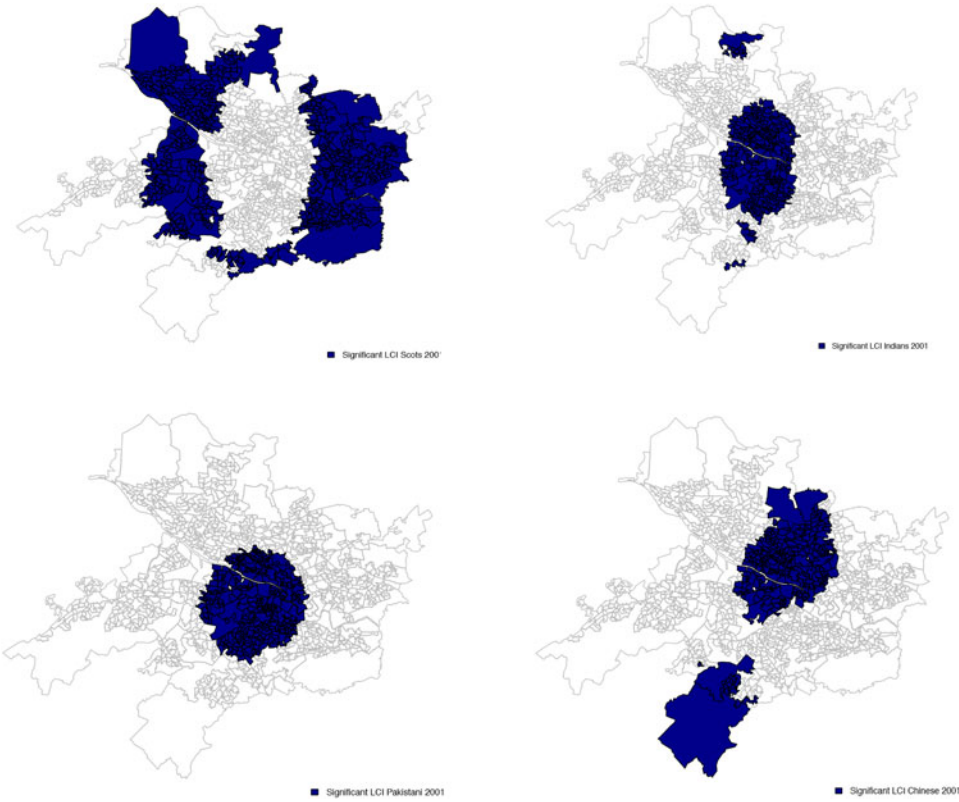


Figure 6. Statistically Significant LCI for 8 km radius, 2001 Census, Datazones. Source: Own calculations based on Scottish Census 2001.

out that proximity to family is a more important factor than ethnic networks and religious and cultural facilities.

The statistically significant LCI's also show that ethnic minority groups are partly clustered in the same neighbourhoods in an around the city centre. These neighbourhoods have relatively high shares of immigrants, and clustering of different ethnic groups. The different ethnic minority groups do not seem to cluster in the same neighbourhoods as the Scottish population. This might indicate a level of segregation.

We now turn to the results from the regression model in which we explain the size of ethnic mover flows in terms of neighbourhood mix and local centralisation indices. The regression includes spatial dummies at the datazone level to control for other local area characteristics that might explain ethnic mover inflows (a datazone comprises between 500 and 1,000 household residents). Table 4 shows that the share (P_{xi}) and LCI (C_{ik}) are both statistically significant for explaining inflows of ethnic groups in neighbourhoods in Glasgow, except for the Scottish house owners. Exposure to the own ethnic group, the number of local possible connections in terms of homophily, is thus positively statistically significant for all groups. However, the central location of a neighbourhood in the network of neighbourhoods is only statistically significant for the ethnic minority groups indicating that mover flows are positive into neighbourhoods that are relative central concentrations of ethnic

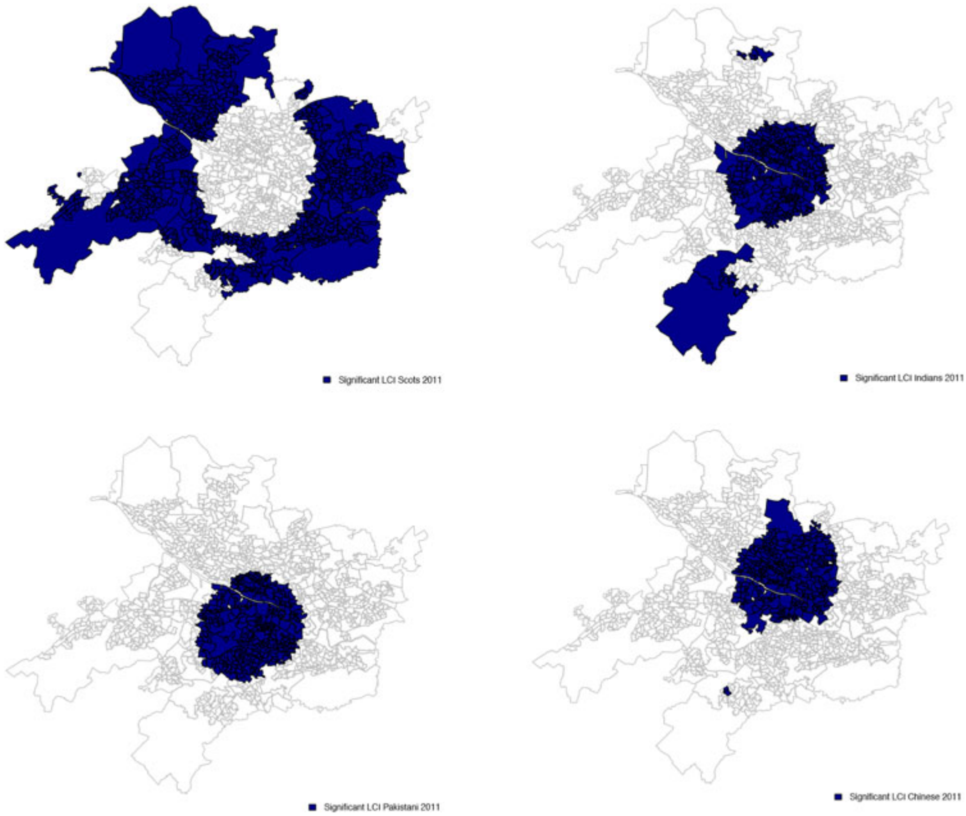


Figure 7. Statistically Significant LCI for 8 km radius, 2011 Census, Datazones. Source: Own calculations based on Scottish Census 2011.

homophily. As the Scottish population is the majority ethnic group in most neighbourhoods, the insignificant effect is not surprising; Scottish residents live in neighbourhoods that are relative central concentrations of the own ethnic group no matter which neighbourhood it is.

The ethnic specific constants indicate that house owner inflows for the ethnic minority groups are in general lower than the inflows of Scottish house owners (the reference group). We also find a negative, statistically significant coefficient for the median house price. Mover flows are smaller into neighbourhoods with higher house prices. This is partly due to the tendency of more expensive houses to have lower owner turn-over rates as well as areas with more expensive houses to have less houses in general, hence a lower potential inflow of people.

To explore the size of the homophily effects, we first interpret the results by calculating the marginal effects at specific points of the distribution of the share and LCI (following Bakens *et al.*, 2018). Table 5 shows the marginal effects at the 5th, 25th, 50th, 75th, and 95th percentiles. Overall, the share of the own ethnic group has a larger effect on the number of inflows than the LCI or the median house price. Marginal changes in house prices hardly decrease the size of the mover flow at any point in the distribution of house prices. The larger the own ethnic share in a datazone, the higher the inflow of that ethnic group. Except for the Scottish house

Table 4. Negative binomial regression results.[a]

	Coefficient	Standard errors
Scottish		
Indian	−2.340***	0.191
Pakistani	−1.816***	0.190
Chinese	−3.368***	0.192
P_{xi}		
Scottish	1.166***	0.210
Indian	16.080***	2.066
Pakistani	5.795***	0.602
Chinese	19.511***	2.482
C_{ik}		
Scottish	−0.179	0.223
Indian	2.400***	0.271
Pakistani	1.878***	0.138
Chinese	2.585***	0.346
Median house price (x £10,000)	−.0010***	0.003
Constant	−1.199***	0.284
Observations	4632	
Pseudo log-likelihood	−9532.7454	
α[b]	0.129***	0.093
Wald test χ^2 (df)[c]	27256.72*** (235)	

Source: Own calculations based on RoS and Scottish census data, 2001.
[a]The statistical significance of coefficients is indicated with ***, **, * for the 0.01, 0.05, and 0.1 significance levels, respectively. The number of houses sold in the period is included in the estimation as exposure variable. Intermediate datazone level dummies are included but not reported here.
[b]The significance of α is based on a χ^2 likelihood-ratio test for overdispersion estimated on a model with nonrobust standard errors with the null hypothesis being that the model is Poisson, corresponding to H0: $\alpha = 1$.
[c]The Wald test performs a full-model test of joint significance of all coefficients, with the null hypothesis being that all coefficients are simultaneously equal to zero.

Table 5. Marginal effects of the negative binomial regression results at percentiles.[a]

	5%	25%	50%	75%	95%
Median house price (x £10,000)	−0.131**	−0.128**	−0.126**	−0.121***	−0.111***
P_{xi}					
Scottish	4.095***	5.100***	5.493***	5.640***	5.755***
Indian	2.506***	2.506***	2.564***	2.870***	4.227***
Pakistani	1.667***	1.667***	1.698***	1.813***	2.851***
Chinese	1.156***	1.156***	1.184***	1.296***	1.800***
C_{ik}					
Scottish	−0.831	−0.803	−0.785	−0.770	−0.752
Indian	0.165***	0.264***	0.406***	0.557***	0.706***
Pakistani	0.235***	0.319***	0.448***	0.784***	1.571***
Chinese	0.082***	0.109***	0.142***	0.224***	0.364***

Source: Own calculations based on RoS and Scottish census data, 2001.
[a]The statistical significance of marginal effects is indicated with ***, **, * for the 0.01, 0.05, and 0.1 significance levels, respectively, and calculated using the variance-matrix of Huber-White robust standard errors from the regression results presented in Table 4. Unit changes in x are 0.1, except for the median house price for which a unit change is £10,000.

owners, we find the same effects for the LCI. However, the effect is much smaller for the LCI than for the shares. For the ethnic groups considered here, we conclude that the homophily horizon is relatively short, indicating that it is the ethnic mix of the immediate locality, rather than the spatial ordering of ethnicity in the surrounding neighbourhoods, that dominates homophily behaviour. Interpreting these results through the lens of social network theory, our findings suggest that the potential for within-neighbourhood connections are more important than the position in the overall network of neighbourhoods. Relative to the effect of the share of the own

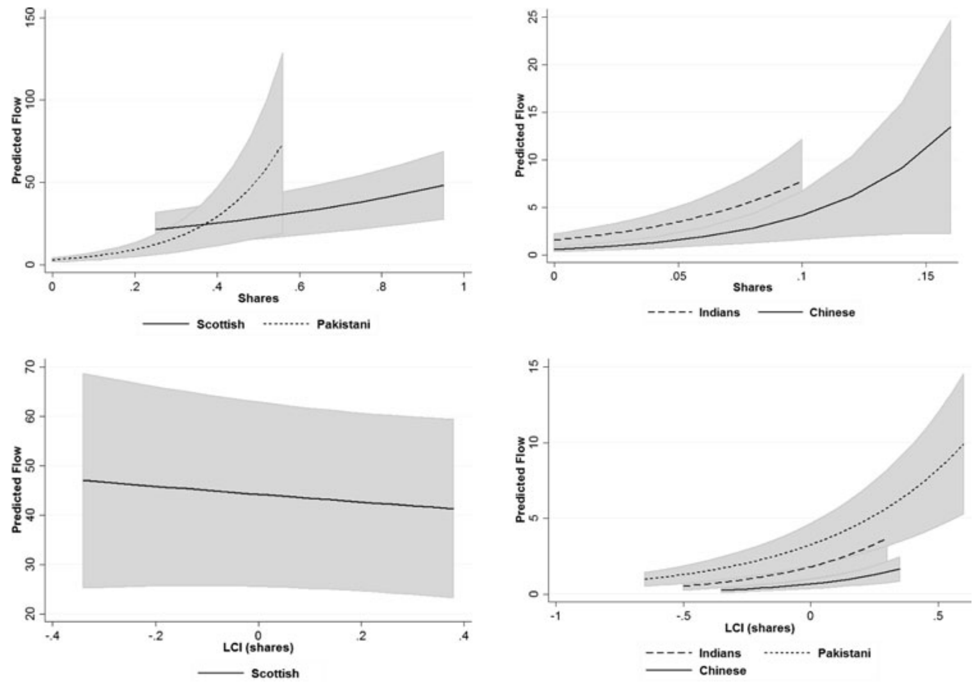


Figure 8. Predicted flows. Source: Own calculations.

ethnic group, it is of less importance for the size of mover flows whether a neighbourhood is the centre of ethnic concentration, or at the boundary of an ethnic concentration. Note, however, that neighbourhoods with relatively high homophily are also at the centre of pockets of statistically significant ethnic concentration, so there is likely to be a degree of conflation of the two effects, particularly for the larger ethnic groups.

The results discussed above are confirmed by the predicted mover flows in Figure 8 which shows the (absolute) predicted size of the inflow for the distribution of the shares and LCI (between 5 and 95 percentiles). Local datazone ethnic shares are more important for the size of mover flows than the spatial population composition of the surroundings of the datazone. Again, for the Scottish house owners the effect of the LCI is not significant. The effect of the own ethnic share and the LCI are nonlinear, and shows that the higher the own ethnic share or the LCI, the higher the increase in the predicted flow.

There is much research on other cities and countries that support the results we find for Glasgow. Bakens *et al.* (2018) find similar effects for the role of homophily, the presence of the own ethnic group, for ethnic groups in the city of Amsterdam and the Hague in the Netherlands. Ibraimovic & Masiero (2014) find homophily in neighbourhood choice in Switzerland, i.e. individuals tend to choose neighbourhoods with a higher concentration of the own ethnic groups, especially lower educated individuals. They point out that this might indicate that the lower educated benefit more from ethnic networks in terms of socio-economic outcomes. Catney (2016a, 2016b) finds that neighbourhoods in England and Wales are becoming more diverse with a growing share of ethnic minorities in the population. Johnston *et al.* (2015)

find the same results for the neighbourhoods in London, but the pattern is rather specific. The neighbourhoods that are predominantly White become more diverse, and the already very ethnically diverse neighbourhoods also become more diverse, but not with Whites. Our results for Glasgow do not necessarily contradict these findings. There are many datazones in the centre of Glasgow that are centres of clusters for more than one ethnicity, indicating that these need to be diverse neighbourhoods. If the minority groups are small, clustering of ethnic groups in neighbourhoods that are at the aggregate diverse is a very likely outcome (see also Bakens *et al*., 2018 for a discussion on these outcomes). We also see a suburbanization of the minority groups, but to different areas than those where the White Scottish people live. This latter result indicates that there are no statistically significant clusters of other ethnic groups in these predominantly White Scottish areas, although there may be an increase in ethnic minorities in these areas. On the basis of this research we thus do not find real evidence for an increase in White Flight. There is ethnic assimilation of already diverse neighbourhoods (i.e. neighbourhoods become more and more diverse), but we do not have conclusive evidence on spatial assimilation in neighbourhoods that are predominantly populated by White Scottish people.

Different papers also point out that some ethnic minority groups are able to move into neighbourhoods with predominantly natives through increasing their socio-economic position, but that the increase of the minority population through natural population growth generally increases segregation (Zwiers *et al*., 2016). As Rosenthal and Ross (2015) show that the socio-economic characteristics of neighbourhoods generally only change very slowly, observed changes in neighbourhood diversity is likely also determined by ethnic minorities changing incomes. For the Netherlands, Bolt *et al*. (2008) find that native Dutch are more likely to move away from neighbourhoods with many ethnic minority groups.

5. Conclusion

We have explored an issue which we believe has been largely overlooked in the voluminous literatures on segregation and ethnic homophily, namely: how near or distant is the homophily horizon in the residential location decisions of ethnic minority groups? In other words, how large an area do movers consider to be relevant when deciding whether the ethnic mix of a locality is sufficiently similar to their own ethnicity? When deciding where to purchase a home, is the homophily horizon a near one, focussed exclusively on the residents of the neighbourhood in question? Or is the homophily horizon one that is distant, reflecting the tendency to consider the ethnic mix not only of the neighbourhood itself but also that of the surrounding ones? The answer, when combined with the predictions of our theoretical model, has profound implications for how potent the overall sorting tendencies of the housing market are in driving the housing market towards segregation. In principle, the more distant the homophily horizon, the more quickly the housing market will tend towards segregation, other things being equal.

These questions presume, of course, that homophily exists and so we first explored whether this was in fact the case. We found that ethnic homophily does indeed characterise the mover flows of homeowners in Glasgow. The presence of people from a specific ethnic group in a neighbourhood has positive and statistically significant effects on the flow of people from that ethnic group moving into that neighbourhood. We find that the spatial ordering of ethnic composition of surrounding neighbourhoods also has a positive and statistically significant effect on the size of mover flows, but with a much smaller magnitude. This suggests a relatively near horizon for homophily effects, which implies less potent tendencies towards segregation in the housing system as a whole in a Schelling view of the world. To our knowledge, this is the first time the notion of homophily horizons has been explored and tested empirically. Interpreting these results from a network perspective we conclude that the position of a neighbourhood in the wider ethnic city network of neighbourhoods is less important than the potential for ethnic connections within the same neighbourhood.

The inference radius of 8 km used to delineate the homophily horizon in Glasgow is much larger than the one computed by Folch & Rey (2015) for Pheonix, Arizona, which may indicate interesting differences in homophily behaviour between the two cities. However, further analysis would be needed to confirm this as there is no estimation in Folch and Rey (nor in any other published study we are aware of) that estimates the impact of their measure on mover flows. For Glasgow, we find that at 8km level, strong spatial patterns emerge and the LCI (local centralisation index) is statistically significant for explaining mover flows. The different ethnic minority groups show display similar patterns of homophily behaviour, so in that sense we find evidence of homophily asymmetry at least at the level of aggregate mover flows between neighbourhoods.

Future work should focus on exploring homophily horizon effects at the individual level. We know from research in the US that there are nuanced aspects to the nature of homophily when considered in terms of individual household preference. For example, even though individuals may have preferences for the ethnic composition of their residential neighbourhood, the preferences between ethnic groups for the ethnic neighbourhood composition may be incompatible (Bayer *et al*., 2014; Vigdor, 2003). Other potentially fruitful avenues of future research include measuring the variation in homophily effects across cities and the potential correlation with social and individual outcomes in terms of well-being, education and employment over the life-course, and how these effects might by mitigated or exacerbated by ethnicity and neighbourhood composition. Also, of potential importance is the impact of homophily horizons on long-run segregation patterns—as illustrated by the simple Schelling model in the Introduction, the effect could be substantial.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was conducted under the auspices of the Location Dynamics, Owner Occupation and Ethnicity in Scotland project, funded by the Economic and Social Research Council (ESRC) as part of the Applied Quantitative Methods Network: Phase II, grant number ES/K006460/1.

Notes on contributors

Jessie Bakens is a researcher and project leader at the Research Centre for Education and the Labour Market (ROA) of the Maastricht University School of Business and Economics.

Gwilym Pryce is a Professor of Urban Economics in the Sheffield Methods Institute and the Department of Urban Studies and Planning, University of Sheffield.

Notes

1. The model is developed in Excel and is available on request.
2. We only have inflows of house owners into neighbourhoods, while Bakens *et al.* (2018) estimate a full gravity model of mover flows between neighbourhoods of Amsterdam and the Hague in the Netherlands. Our estimation differs from Bakens *et al.* (2018) in that we focus on homophily horizons.
3. In this article we use the term 'ethnic' groups to refer to the ethnic descent of an individual based on the cultural, ethnic, and linguistic analysis of one's name.
4. Because we estimate a partial gravity model, we do not know which neighbourhood the house owners left. It is therefore difficult to draw conclusions of whether this results in more, equal, or less segregation in Glasgow.
5. Our unit of analysis is the datazone derived from the Scottish census geographies. A datazone consists of about 500 to 1,000 households. We use datazones and the term neighbourhoods interchangeably throughout the text.
6. Although the transaction data lacks information on the dwelling characteristics, the exact location of the dwelling provides options to link the data to detailed neighbourhood characteristics from Scottish Neighbourhood Statistics, including the type of dwellings.
7. ONOMAP is the software based on the identification strategy of names built by the Department of Geography at University College London. See for further information on the construction of the underlying database Mateos *et al.* (2007). After finishing the analysis with the ONOMAP software, the name of the software has changed into Onolytics, See https://onolytics.com/ for more information.
8. A detailed description of the matching process and the validation of the name-matching with ethnicity from the official census data is available upon request.
9. The urban TTWA consists of 1169 datazones, of which 11 do not have information on population or house prices. These datazones are excluded from the analysis.

References

Alba, R.D., Logan, J. R., Stults, B. J., Marzan, G. & Zhang, W. (1999) Immigrant groups in the suburbs: A reexamination of suburbanization and spatial assimilation, *American Sociological Review*, 64(3), pp. 446–460.

Bakens, J., Mulder, P. & Florax, R. J. G. M. (2018) Ethnic drift and white flight: A gravity model of neighborhood formation, *Journal of Regional Science*. doi:10.1111/jors.12390.

Bayer, P., Fang, H. & McMillan, R. (2014) Separate when equal? Racial inequality and residential segregation, *Journal of Urban Economics*, 82(1), pp. 32–48.

Bayer, P., McMillan, R. & Rueben, K. S. (2004) What drives racial segregation? New evidence using census microdata, *Journal of Urban Economics*, 56(3), pp. 514–535

Becker, G. S. & Murphy, K. M. (2000). *Social Economics: Market Behavior in a Social Environment*. Cambridge and London: The Belknap Press of Harvard University Press.

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(7), pp. 1359–1384.

Borjas, G. (1995) Ethnicity, neighborhoods, and human capital externalities, *American Economic Review*, 85(3), pp. 365–390.

Card, D., Mas, A. & Rohstein, J. (2008) Tipping and the dynamics of segregation, *Quarterly Journal of Economics*, 123(1), pp. 177–218.

Catney, G. (2016a) Exploring a decade of small area ethnic (de-)segregation in England and Wales, *Urban Studies*, 53(8), pp. 1691–1709.

Catney, G. (2016b) The changing geographies of ethnic diversity in England and Wales, 1991–2011, *Population, Space and Place*, 22(8), pp. 750–765.

Dustmann, C. & Fabbri, F. (2003) Language proficiency and labour market performance of immigrants in the UK, *The Economic Journal*, 113(484), pp. 695–717.

Folch, D. C. (2012) *The centralization index as a measure of local spatial segregation*. PhD Thesis. Tempe, AZ: Arizona State University.

Folch, D. C. & Rey, S. J. (2015) The centralization index: A measure of local spatial segregation, *Papers in Regional Science*, 94(3), pp. 1–22.

Giuffre, K. (2013) *Communities and Networks: Using Social Network Analysis to Rethink Urban and Community Studies*, Kindle ed. (Cambridge: Polity Press, Wiley).

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

Ibraimovic, T. & Masiero, L. (2014) Do birds of a feather flock together? The impacts of ethnic segregation preferences on neighborhood choice, *Urban Studies*, 51(4), pp. 693–711.

Ioannides, Y. M. & Zabel, J. E. (2008) Interactions, neighborhood selection and housing demand, *Journal of Urban Economics*, 63(1), pp. 229–252.

Johnston, R., Poulson, M., & Forrest, J. (2015) Increasing diversity within increasing diversity: The changing ethnic composition of London's neighborhoods, 2001–2011, *Population, Space and Place*, 21(1), pp. 38–53.

Krysan, M. & Farley, R. (2002) The residential preferences of blacks: Do they explain persistent segregation?, *Social Forces*, 80(3), pp. 937–980.

Kuminoff, N. V., Smith, V. K., & Timmins, C. (2013) The new economics of equilibrium sorting and policy evaluation using housing markets, *Journal of Economic Literature*, 51(4), pp. 1007–1062

Long, J. S. (1997) *Regression Models for Categorical and Limited Dependent Variables* (Thousand Oaks, CA: Sage Publications).

Massey, D. S. & Denton, N. A. (1985) Spatial assimilation as a socioeconomic outcome, *American Sociological Review*, 50(1), pp. 94–106.

Massey, D. S. & Denton, N. A. (1988) Suburbanization and segregation in US metropolitan areas, *American Journal of Sociology*, 94(3), pp. 592–626.

Massey, D. S. & Tannen, J. (2017) Suburbanization and segregation in the United States: 1970–2010, *Ethnic and Racial Studies*, 41(9), 1594–1611. doi: 10.1080/01419870.2017.1312010

Mateos, P. (2014) *Names, Ethnicity and Populations: Tracing Identity in Space* (Berlin Heidelberg: Springer Verlag).

Mateos, P., Webber, R. & Longley, P. (2007) *The Cultural, Ethnic and Linguistic Classification of Populations and Neighbourhoods using Personal Names* Working Paper 116. Seattle, WA: CASA.

McGarrigle, J. L. (2010) *Understanding Processes of Ethnic Concentration and Dispersal:* South Asian Residential Preferences in Glasgow IMISCOE Dissertations. Amsterdam: University Press.

McPherson, M., Smith-Lovin, L. & Cook, J. M. (2001) Birds of a feather: Homophily in social networks, *Annual Review of Sociology*, 27, pp. 415–444.

Nandi, A. & Luthra, R. R. (2016) *Who Experiences Ethnic and Racial Harassment? Written Evidence Submitted by the Institute for Social and Economic Research, University of Essex. Home Affairs Committee. Hate crime and its Violent Consequences Inquiry, HCR0090* (Colchester: University of Essex: Institute for Social and Economic Research).

Portes, A. & Sensenbrenner, J. (1993) Embeddedness and immigration: notes on the social determinants of economic action, *American Journal of Sociology*, 98(6), pp. 1320–1350.
Rosenthal, S. S. & Ross, S. L. (2015) Change and persistence in the economic status of neighborhoods and cities, in: G. Duranton, V. Henderson, & W. Strange (Eds) *The Handbook of Regional and Urban Economics*, pp. 1047–1120 (North Holland: Elsevier Science).
Rytina, S. & Morgan, D. L. (1982) The arithmetic of social relations: The interplay of category and network. *American Journal of Sociology*, 88, pp. 88–113.
Santos Silva, J. M. C. & Tenreyro, S. (2006) The log of gravity, *Review of Economics and Statistics*, 88(4), pp. 641–658.
Schelling, T. C. (1971) Dynamic models of segregation, *Journal of Mathematical Sociology*, 1(2), pp. 143–186.
Vigdor, J. L. (2003) Residential segregation and preference misalignment, *Journal of Urban Economics*, 54(3), pp. 587–609.
Waldfogel, J. (2008) The median voter and the median consumer: Local private goods and population composition, *Journal of Urban Economics*, 62(2), pp. 567–582.
Winkelmann, R. (2008) *Econometric Analysis of Count Data, 5th ed.* (Berlin, Heidelberg: Springer-Verlag).
Zwiers, M, van Ham, M. & Manely, D. (2016) *Trajectories of Neighborhood Change: Spatial Patterns of Increasing Ethnic Diversity* Discussion Paper no. 10216. Bonn, Germany: IZA.