Environment and Planning B:
Urban Analytics and City Science

The Conflicting Geographies of Social Frontiers: Exploring the asymmetric impacts of social frontiers on household mobility in Rotterdam

Journal:	*Environment and Planning B: Urban Analytics and City Science*
Manuscript ID	EPB-2022-0253.R1
Manuscript Type:	Manuscript
Keywords:	social froniters, segregation, geographical mobility, Netherlands, immigration

The Conflicting Geographies of Social Frontiers

Exploring the asymmetric impacts of social frontiers on household mobility in Rotterdam

Abstract: Social frontiers arise when there are sharp differences in the demographic composition of adjacent communities. This paper provides the first quantitative study of their impact on household mobility. We hypothesise that conflicting forces of white flight and territorial allegiance lead to asymmetrical effects, impacting residents on one side of the frontier more than the other due to differences in the range of housing options available to different groups, and different symbolic interpretations of the frontier. Using Dutch registry data for the city of Rotterdam we identify ethnic social frontier locations using a Bayesian spatial model (Dean et al 2019), exploiting the data's one hundred metre resolution to estimate frontiers at a very small spatial scale. regression analysis of moving decisions finds that the ethnic asymmetry of the frontier matters more than ethnicity of individual households. On the ethnic minority side of the frontier, households of all ethnicities in the 28-37 age range have reduced probability of moving compared to non-frontier parts of the city. The opposite is true on the Dutch native side of the frontier. We supplement this analysis with flow models which again find strong frontier effects. Our findings illustrate how the study of social frontiers can shed light on local population dynamics and neighbourhood change.

1. Introduction

Turner's (1893) frontier thesis has shaped the etymology and symbolic potency of the "frontier". As a result, frontiers evoke a wider set of meanings than merely bureaucratic delineation. They denote zones of contested space that drive, and are driven by, the aspirations and movements of people. Frontiers are to be advanced, defended, valorised and feared. They engender a sense ingroup

purpose and solidarity. And whilst it is true that frontiers represent meeting of two contrasting communities, the term is not usually suggestive of integration, mixing or conviviality. Rather, frontiers carry overtones of contrast, separation and latent tension (Iyer and Pryce, 2022).

Mindful of these historical and etymological connotations, researchers have recently adopted the term "social frontier" to describe a particular form of residential segregation where "steep differences in the racial, ethnic, religious, cultural or social characteristics" (Dean et al. 2019, p.271) occur at the border between adjacent communities. However, as with the broader use of the word, social frontiers are potentially more than purely geographical phenomena. They can take on symbolic meanings that influence attitudes, behaviour and wellbeing (Piekut and Pryce, 2022; Iyer and Pryce, 2022).

The emerging interest in social frontiers as a new focal point for segregation research is, in part, a response to longstanding concerns that the literature "does not address what happens at places where groups border" (Legewie and Schaeffer, 2016, p.131). Sack (1983, p.55), for example, lamented the tendency of "conventional spatial analysis" to overlook territoriality. As a result, "many of the forces moulding human spatial organisation" remain unexamined (ibid). Similar criticisms have been made by Spielman and Logan (2013), Kramer (2017) and Dean et al. (2019). Whilst there have been inroads into the quantitative investigation of social frontier impacts on crime and anti-social behaviour (e.g. Legewie and Schaeffer 2016; Dean et al. 2019), the wide

spectrum of other potential effects--such as psychological, health, demographic, economic and social impacts--remain largely unexplored.

The contribution of the current paper is to provide the first exploration of the relationship between social frontiers and residential mobility. We start by providing the first estimates of social frontier location in Rotterdam (and indeed the Netherlands) using the empirical approach developed by Dean et al. (2019). This adopts Lee and Mitchell's (2013) locally adaptive spatial conditional autoregressive Bayesian estimation method to identify step changes across adjacent neighbourhoods in the proportion of non-Western migrants. We then link these social frontier estimates (see map in Supplementary Material) to household moving behaviour to explore two key questions: (i) the specific question of how household moving decisions vary by proximity to social frontiers; and (ii) the broader question of whether social frontiers are associated with asymmetric impacts. With respect to the first, whilst there are powerful reasons to expect social frontiers to have significant impacts on geographical mobility (as we discuss below), to our knowledge there have been no attempts to quantify them. Indeed, the role of social frontiers seems to have been comprehensively overlooked which is surprising given the voluminous literature on household mobility (see review by Dieleman 2001). The second question has wider implications in the sense that the many and varied theoretical impacts of social frontiers (Iyer and Pryce 2022) could radically differ depending on which side of the frontier one is on. Again, to our knowledge, none of the existing papers on social frontiers, including those looking at crime impacts (e.g. Dean et al. 2019), have considered the possibility of asymmetric effects.

We conceptualise social frontiers as the embodiment of contradictory geographies: where conflicting forces collide and yield asymmetric impacts. Forces that encourage some households to stay are counterbalanced by factors that motivate others to leave. We do not attempt to disentangle these causal processes. Rather, we present a detailed description of the net effect of these processes with respect to the moving decisions of native-born households versus nonwestern migrants.

The remainder of this paper is structured as follows. In section 2 we review the literature on geographical mobility and social frontiers. Having highlighted the relevant gaps, we set out our main hypotheses in section 3. These form the basis of our empirical model. In section 4, we describe our methods, data and socio-geographic context. This is followed with a summary of our key results in section 5 followed by a discussion of our findings in section 6. We conclude in section 7 with a brief summary and suggestions for further research.

2. Literature

Social frontiers arise when there is minimal residential mixing at the border between two adjacent but markedly different communities (Dean et al., 2019). The attribute of difference could be country of birth (as in the present study), or it could be ethnicity, religion or some other social, cultural or economic characteristic. Although social frontiers may sometimes form around pre-existing physical barriers such as roads and railway tracks (Noonan 2005), this is not an

essential feature. The social frontiers between Catholics and Protestants in Belfast, for example, predated the erection of peace-walls.

Early 'border area' theoretical literature focussed on how social frontiers emerge as the product of individual preferences and the free movement of households, for example (Yinger 1976). Recent theoretical work (Dean et al 2019; Iyer and Pryce 2022) has focussed on potential impacts of frontiers on community cohesion and crime. Compared with more granular borders, frontiers reduce the potential for positive contact between groups, leading to worsening intergroup relations (Allport 1954). Social frontiers may also take on territorial meaning (Sack 1983), evoking defensive and hostile behaviour (Dean et al 2019; Iyer and Pryce 2022).

Work on the impact of social frontiers has explored the effect on mental health (MaGuire et al. 2016), anti-social behaviour (Legewie and Schaeffer 2016) and crime (Dean et al. 2019). MaGuire et al. (2016) found that proximity to peace walls in Belfast were associated with substantially higher rates of take-up of anti-anxiety and antidepressant drugs. Dean et al. (2019) found higher rates of crime near social frontiers in Sheffield. Legewie and Schaeffer (2016) also reported higher rates of anti-social behaviour near social frontiers, though the effect was lower for the most pronounced frontiers.

As far as we are aware, however, there have been no empirical studies quantifying the relationship between social frontiers and geographical mobility. Iyer and Pryce (2022) provide theoretical reasons for expecting lower rates of geographical mobility near the frontier but do not attempt empirical verification.

There is also a longstanding finding that rates of out-migration are often lower for households from a specific ethnic group if they are surrounded by other households from the same ethnicity. The higher the proportion of own-ethnicity, the lower the propensity to move (e.g. Clark, 1989, p. 171). This effect may arise from the various benefits to migrant households of living near those from a similar background; such as protection from racial harassment, preservation of cultural and religious identity, greater dependence on local social networks for employment opportunities, and the need for social support and advice, especially where there are language barriers. The fewer the number of alternative neighbourhoods offering such benefits, the less likely a resident will be tempted to move away once housed there. In the Netherlands, for example, the closer immigrant groups are to native Dutch culture, the more likely they are to move there in the first place, and the lower the probability that they will move away (Zorlu and Mulder, 2008, p. 248).

As noted in the Introduction, social frontiers may represent zones of contradictory geographies where potent pull-factors that persuade households to stay, operate alongside push-factors encouraging them to move. Residential mobility on frontiers is determined by the balance of these forces and how their strength changes over geographical distance (cf. the concept of "homophily horizons" proposed by Bakens and Pryce, 2019) and group characteristics. Social frontiers may also introduce their own unique dynamics, created by the type of encounter they form between particular socio-ethnic groups. For example, social frontiers may evoke territorial behaviour, enhancing neighbourhood allegiance and a sense of obligation to 'defend' the frontier. This may reduce geographical

mobility and restrict educational and labour market progression (Iyer and Pryce, 2022).

'Cliff-edge' social frontiers are, in one sense, the strongest separation of ethnic groups it is possible to create. T wo ethnicities are divided as if by a wall, whether or not there actually is one. Relative to neighbourhoods where residential mix at the border creates a more gradual transition between communities , social frontiers offer far fewer opportunities for positive contact (Allport, 1954). On the other hand, compared to living at the core of mono-ethnic areas, social frontiers represent an interface zone with the potential to facilitate inter-ethnic contact, depending on the physical structure of the frontier itself and flows across it (for example, retail, commuting or school flows).

So when considering the 'housing bundle' – all the factors affecting the household's decision to move – social frontiers may constitute an additional dimension that has the potential to transform existing elements of the bundle, especially for particular stages of the life-cycle (Dieleman, 2001 p.250, 254). It is also important to note that migrant households, particularly younger ones, may have markedly different decision sets (Zorlu and Mulder, p.250). As our contradictory geographies hypothesis suggests, we must consider how the social-frontier-specific bundle could vary in polarity and intensity across the frontier. There are two key aspects to consider. First, preferences are heterogeneous across different ethnic groups and – perhaps more importantly – within them (Clark, 1991, p. 2). Second, when considering hypotheses of social frontier

effects, it is vital to consider that 'revealed preference' of location choice may not capture actual preference (Ibraimovic and Hess, 2017).

Considering the first point, suppose all household ethnic neighbour preferences can be placed into three possible categories: homophily, indifference and heterophily (diversity-seeking). The first preference category, homophily, denotes a preference to locate near households of the same group. This is the assumed outlook of households in the celebrated Schelling (1978) model which shows how even relatively weak homophily tendencies can lead to extreme segregation. An often overlooked aspect of this model is how it leads not only to the separation of communities but also to abrupt rather than gradual residential transitions between them. Perhaps one of the most important yet unexplored questions raised by the Schelling model, therefore, is not why social frontiers sometimes arise, but why they are not ubiquitous (Piekut and Pryce, 2022).

One answer is that households may fall into the indifference preference category: they are ambivalent to the ethnic composition of neighbourhoods when making location choices. If all households were indifferent, social frontiers would not arise, but neither would segregation (other things being equal) because households would have no aversion to being in the minority. The same would be true for our third category of preference types: if households are heterophilous (i.e. have a preference for diversity), they will actively seek residential mix, ruling out both segregation and social frontiers.

So why, then, might we observe segregation without social frontiers – adjacent communities with blurred rather than abrupt borders? One answer is

that other drivers of location choice (e.g. the desire to live close to employment and amenities such as schools, transport, and leisure) mask 'pure' preferences with regard to social mix. Another explanation is that preferences are not uniform. If we relax the assumption of homogeneous preferences then we would expect those who prefer (or are indifferent to) diversity to be sorted by the market to live at the interface between the two communities, creating blurred boundaries. Thus, the degree to which we see residential mixing at the border as opposed to social frontiers may reflect the heterogeneity of preferences within each community (Piekut and Pryce, 2022). The prevalence of social frontiers, therefore, may indicate relatively homogenous preferences within each group; social frontiers are more likely to emerge when there is a shortage of residents who are ambivalent or heterophilous with respect to ingroup vs outgroup residential mix.

Constrained Choice

Frontiers might also arise because unconstrained households who do not wish to be there move away, leaving behind others with more constrained moving options (e.g. those reliant on social housing). Ibraimovic and Hess (2017), for example, using a 'stated preference' approach, find that preferences are not uniform. While we cannot access stated preference with our Dutch registry data, it is possible to use proxies to inform discussion on what role constraints could be playing – in particular, whether there are differences in

household wealth and tenure on frontiers that, all other things equal, would indicate more constrained options. We do this in the discussion of the results.

White Flight

Social frontiers may elicit responses from native incumbents similar to those described in the 'white flight' literature . Easton and Pryce (2018), for example, note that the motivation of the majority-group out-movers may not be homophily (the tendency to connect with those most like oneself) but heterophobia—the fear or dislike of people socially "other" to oneself. Territorial attitudes may be heightened if social frontiers emerge as the majority group "retreat" from areas they have long occupied . A common weakness of research on 'white flight' is the tendency to rely on aggregate data available at aerial unit levels that are unhelpfully large for studying this phenomenon. Aerial unit size is crucial because awareness of ethnic minority in-movers will likely decay rapidly with distance (Easton and Pryce 2020). Analysis at census tract level, for example, potentially misses important variation at smaller spatial scales and so more attention needs to be paid to moving behaviour at the micro-level, particularly the dynamics of micro-neighbourhoods surrounding social frontiers. To our knowledge, there has been no such study to date.

3. Hypotheses

Our review of the literature leads to three testable hypotheses:

H1: Ethnic Density Mobility Effects*: Households will be less likely to move away from a neighbourhood with a high-proportion of their own group; this tendency will likely be heightened or ameliorated through individual characteristics (such as age and social class), and the wider social and demographic status of the group.*

Because social frontiers inevitably entail high proportions of a particular ethnic/social group either side of the frontier, there will always be the potential for an ethnic density effect near social frontiers. This is likely to be especially true of ethnic groups most dependent on local own-group social networks. Cultural and socio-economic vulnerability will likely heighten this effect.

H2: Constrained Choice Dampens Mobility: *households from minority or low income groups may be less likely to move away from social frontiers due to limited alternative residential options relative to the majority native group.*

This raises the possibility of asymmetric mobility effects if a community on one side of the frontier has less constrained choice sets. Whilst this may be less true in Rotterdam, where the native / non-native split is so close to fifty

percent, limited financial resources among nonwestern migrants may nevertheless constrain their relocation options.

H3: Neighbourhood allegiance: *Social frontiers have the potential to elicit territorial behaviour, imbuing households with a sense of obligation to remain in the neighbourhood, reducing social mobility.*

A social frontier may represent a zone of threatened identity where residents become preoccupied with countering the psychological threat. In this context, clearly demarcated boundaries can take on symbolic meaning (Piekut and Pryce 2022) and elicit territorial behaviour such as a stronger sense of allegiance to the neighbourhood. One way of viewing this hypothesis is in terms of it heightening the ethnic density effect. In other words, we would expect the impacts of ethnic density to be higher close to social frontiers .

These three hypotheses will likely have different impacts for different population groups depending on their life-stage, financial resources, overall numerical prevalence in the city/country and perceived sense of socio-economic and cultural vulnerability. The corollary is that when social frontiers divide groups with different combinations of these attributes, the frontier will generate asymmetrical impacts on household mobility.

4. Methods and Data

Frontier detection and stable frontiers

We estimated frontiers between native Dutch households and nonwestern migrant households using a spatial Bayesian model which returns a probability distribution describing the likelihood that two geographical zones have a frontier between them (Dean et al., 2019). These probability values are then normalised (we label these Φ). The frontiers used in the mobility model are stable frontiers that meet the following double-hurdle criteria: (a) *they are statistically significant* in the Bayesian frontier detection algorithm for the five year period in which we measure mobility (2012-17); and they are also (b) *substantively significant*, defined as having a value of Φ above threshold of one standard deviation from the mean. Extensive sensitivity testing (see Supplementary Material) suggests that our results are generally robust to the Φ threshold chosen.

In this paper we focus on the relationship between mover dynamics and stable frontiers. We focus on stable frontiers because it makes it more likely that we are observing a genuine social frontier, rather than a temporary spatial discontinuity arising from random population churn, especially prevalent when using such fine-grained data (both spatially and temporally). We also think that the effects of a spatial discontinuity are likely to be more pronounced if the frontier is persistent rather than fleeting . For example, the territorial symbolism of frontiers is likely to be something that emerges as frontiers become embedded . Stable frontiers will also restrict opportunities for intergroup contact over a

prolonged duration, increasing the likelihood of distrust and prejudice emerging between groups (cf Allport's 1954 'contact hypothesis').

Context, data and variables

Geographical Context

The city of Rotterdam is the second largest city of the Netherlands with more than 600,000 inhabitants within the city boundaries. Due to the presence of the port – one of the largest in the world – and related (heavy) industry, the city has a reputation of being a working class city. However, in the recent decade the service sector has become increasingly important (Binnenlandsbestuur, 2020).

The Netherlands has a relatively recent history of large-scale migration compared to other European countries like the UK. Large scale immigration started in the 1960s and 1970s when the Netherlands welcomed large groups of guest workers from mostly Morocco and Turkey. Many of these guest workers stayed in the country and through family reunification the immigrant groups grew rapidly. The city of Rotterdam now has the highest proportion of people with a migration background of any city in the Netherlands. In 2021 the proportion of people with a migration background exceeded that of the native Dutch population; 53% of the population now has an immigration background, against just over 24% for the country as a whole (Onderzoek 2022).

Geographically the city of Rotterdam is divided in two by the river, which also forms an important socio-economic and ethnic divide. The area north of the city is considerably wealthier than the area south of the river. And the poorer southern part of the city is also characterised by a stronger presence of people with a migration background. (See further background information in Supplementary Materials).

Data

Our analysis employs Dutch registry data from Statistics Netherlands. This is a very rich dataset containing details of every Dutch person and every household for every year from 1999 onwards including a unique identification code for each household and person and a link giving the list of individuals that make up every household. The city is divided up into 100x100m grid squares within which we can identify the location of each person and household tracked over time and space. The very large size of the dataset allows the model to be disaggregated into quite small subgroups while still producing statistically robust results.

The outcome variable is a binary flag marking whether each household moved or not between 2012 and 2017 . Whether or not a household stays is deduced by tracing whether a household's ID was in the same 100m grid square in both of those years. Mixed-ethnicity households can have a large range of different origins and make-ups; using just single-ethnicity households allows for better clarity when interpreting the results. In total, these choices provide the

model with over 178,000 households from the Rotterdam urban region (see Supplementary Material for descriptive statistics and further information on the sample).

Ethnicity in Dutch microdata is defined objectively by the country of birth of a person's parents (in contrast to, for example, UK Census data that uses self-reported ethnicity). Any individual born outside the Netherlands, or any individual with either parent born outside of the Netherlands is labelled 'non-native' (either "Western non-native" or "nonwestern non-native"); consequently, a 'native' is at least third generation Dutch. "Western non-native" individuals and households (generally from Indonesia or Japan) have characteristics closer to Dutch natives; in order to maximise the difference across frontiers, we use only "native" and "nonwestern non-native" categories.

The age profile of households is a potentially important factor in moving decisions. Figure 1 shows the age profile specifically for Rotterdam broken down by ethnicity (using the average age of adults in each household). The age profile for nonwesterners skews earlier and has far fewer older and retired households. Movement behaviour varies greatly between age groups; Figure 2 shows this variation for a series of five-year age bands for each ethnic group for the time period 2012-2017. For both ethnicities, the older the household, the more likely they are to have stayed in that five year period. But ethnic differences are striking. For households with an average adult age below the age of 37, native households are relatively *less* likely to have stayed. The opposite is true for households with an average adult age over 37. Because age effects

interact with virtually all variables of interest in the model, we run separate models for three different age bands. Along with the ethnicity model breakdown, this gives six separate logit models (see Figure 3 for headline results and Supplementary Materials for the regression tables).

<Figure 1 here >

<Figure 2 here >

Model strategy

This section explains the two main model approaches used in the paper: first, a series of 6 regressions modelling the probability of households moving and how this differs on and off frontiers (logit versions of the models are presented in the Supplementary Material); and second, a pair of regressions examining person flow behaviour into and out of zones on the frontier. Both model types consider not just whether a household or person is on a frontier, but what side of the frontier they are on. For any given pair of zones on each frontier side, one will have a higher proportion of natives than the other. This side is labelled the "native side" of the frontier; the other the "nonwestern side".

Probability of households moving: frontier vs. non-frontier areas

We estimate six linear probability “mover” regressions where the dependent variable is coded 1 if the household moves, and 0 if it stays. We use

these models to estimate the key determinants households moving in the period 2012-2017. Using a dummy for whether a household resides on a frontier—that is, whether the 100x100m grid square they reside in has a social frontier directly on one of its edges—and whether that frontier is native versus nonwestern, we ask: is the probability of households having moved significantly different in frontier locations compared to non-frontier parts of the city? And are there asymmetric effects – that is, do those moving probabilities differ depending on the side of the frontier the household is on?

We run six separate models, one for each of our three defined age groups per ethnicity where ethnicity of the household is defined as being of native or non-native nonwestern origin, and age bands are defined in terms of the average adult age in household (19 to 27; 28 to 37; 38 to 65). Other key explanatory variables include *income*, measured using a binary variable capturing whether or not a household is above median income in the city; and *ethnic density*, measured as the proportion of population in the household's grid square that is of the same ethnic group.

Income is a central explanatory variable especially for capturing how constrained choice interacts with neighbourhood allegiance and any other social frontier effect. In the results discussion, we dig deeper into income on frontiers to help understand the dynamics taking place there. The own-group ethnic density variable helps us disentangle the well-known homophily effects from the impact of the frontier.

Control Variables

In order to focus on the three hypotheses noted above, we include the following control variables:

- *Tenure*: rented or owner-occupied;
- *Household structure*: single person; couple without children; couple with children; other;
- *Number of children.*

2. Flow models of individual movement behaviour: difference on frontiers compared to non-frontier areas of the city

The regression models only consider whether households moved *away* during the study period. But there is a marked difference in outcome if, for example, households of the same ethnicity are replacing those that move, or if households are moving in, but not out. To address this issue, we also model *flow behaviour*: the balance and scale of movement on frontiers. The unit of observation is grid squares within the city, for which we find a count of the flow of individuals into and out of those grid squares and two control variables: average standardised income and percent owner-occupied in each geographical zone.

We use two flow measures: *migration efficiency* and *turnover* (see e.g. Dennett and Stillwell, 2008). Migration efficiency measures the *polarity of flows*

into and out of a zone. Minus 1 indicates all flows were out, none came in; +1 indicates all flows were in, none left. Zero indicates a balance of inward and outward flows. Migration efficiency provides a measure of polarity but not scale; so we use *turnover* for the latter: the sum of inward and outward flows divided by total zone population.

See Supplementary Material for further details on the method and results.

5. Results

Do social frontiers affect household moving propensities?

Figure 3 provides a summary of the separate model runs for each of the three average-adult-age groups (x axis) and two ethnicities (separate graphs for each ethnicity, and whether that household is on the native or nonwestern frontier side). The estimates show how household moving probabilities differ on frontiers compared to non-frontier areas of the city (y axis).

There is a clear asymmetry: on the native frontier side (left-hand column of graphs), results are either non-significant (error bars cross zero) or moving probabilities are higher compared to similar non-frontier areas of the city. On the nonwestern frontier side (right-hand column), where results are significant, moving probabilities are lower compared to similar non-frontier city areas. The pattern is more consistent on this frontier side for both ethnicities, with both the older age groups showing significantly lower probabilities of moving.

It is worth focusing on how results for the 28 to 37 year old age group on each side of frontiers differ, as these are all significant: for this age range, households of both ethnicities on the nonwestern side of the frontier have lower probabilities of moving - for both ethnicities, compared to non-frontier parts of the city. On the native side, the result is the opposite: households - again of both ethnicities - have higher chances of moving, relative to the rest of the city. This supports the idea that frontier side matters more than household ethnicity for moving behaviour.

The other age groups either confirm this pattern - lower moving probabilities on the nonwestern frontier side for both ethnicities, higher moving probabilities on the native side - or the results are not significant (confidence intervals cross zero).

Sensitivity analysis confirmed that our findings are robust to various changes in model specification (see Supplementary Material).

<Figure 3 here >

Impact of Frontiers on population dynamics: Flow model results

We find evidence for frontier effects on migration efficiency but they appear to depend more on age-group than ethnicity. Nevertheless, we again find clear evidence of asymmetry. For 19 to 27 year old nonwesterners on the nonwestern frontier side, more leave than come in, relative to the rest of the city

(migration efficiency is lower on frontiers; difference of frontier versus non-frontier = -0.057 [95% CI -0.099, -0.014]). For 28 to 37 year olds, the opposite is true: natives on the nonwestern frontier side, more come in than leave, relatively (migration efficiency is higher on frontiers; difference of frontier versus non-frontier = 0.08 [95% CI 0.009, 0.15]).

Turnover models show similar outcomes: turnover on frontiers is *lower* than in other areas of the city; frontiers are more stable in terms of throughput of people. This is true for ethnicities aged 28-37 on the nonwestern frontier side (native turnover difference for frontiers vs non-frontiers = -0.085 [95% CI -0.144, -0.026]; nonwestern turnover difference = -0.105 [95% CI -0.145, -0.065]). It is also true for 38-65 year old nonwesterners on the nonwestern frontier side (difference = -0.06 [95% CI -0.093, -0.028]). And when all ages are combined, the result is the same for nonwesterners on both sides of the frontier (difference on nonwesterners' own side = -0.084 [95% CI -0.118, -0.05]; difference for nonwesterners on native side = -0.09 [95% CI -0.142, -0.036]).

Turnover is generally lower on the nonwestern frontier side, wherever there is a significant result. It is also lower for nonwesterners of all ages on the *native* side. But – it is *higher* for the two younger *native* age groups, on the native frontier side. This highlights the turnover difference for native versus nonwestern people. More details of the flow model results are provided in Supplementary Material.

6. Discussion

We now interpret these results through the prism of our three hypotheses.

Ethnic density mobility effects

For household moving probabilities, the *side* of the social frontier has a stronger effect than the ethnicity match of the household and the grid square it resides in. Consider the nonwestern side of frontiers: both nonwestern and native households are less likely to move, compared to the rest of the city. If homophily forces brought about by surrounding same-ethnicity density were paramount and people were able to exercise their preferences, native households would be more likely to move – including more likely, relative to nonwestern households on the same side.

Migration efficiency results suggest the difference between age groups is important: two of the significant results are on the nonwestern frontier side, but show opposite outcomes for 19-27 year old nonwesterners (more leaving than coming in relative to the city) versus 28-37 year old natives on that side (more coming in than leaving, relatively). This supports the idea that homophily forces are not the strongest factor affecting moving behaviour on frontiers.

Evidence in the opposite direction comes from the result for nonwesterners of all age groups on the native frontier side: more leave than come in, relatively. This goes with the grain of expectations – that the native frontier side would see nonwestern residents' homophily preferences leading them to move away more.

But the overall picture for homophily seems to be one of variety across age groups – and especially across frontier sides. Each side has its own unique dynamic: not just opposite effects, but quite separate sets of effects. This reinforces the idea of frontiers as contradictory geographies: the array of homophily forces there are so varied, it is perhaps not surprising they balance in complex ways.

Some natives on the nonwestern side have lower turnover while some natives on the native frontier side have *higher* turnover. One possible explanation for this pattern is preference heterogeneity *within* ethnic groups leading to sorting – natives on the nonwestern side may be happy to be there, while natives on the native side may be less happy being right next to a nonwestern area.

In terms of homophily, then, this result again supports the idea of homophily heterogeneity and sorting taking place on frontiers. This leaves the question of whether frontiers are generally residentially non-desirable – the following two sections help shed some light on this aspect of homophily.

Constrained choice

Choice constraints and homophily are closely connected: the homophily preferences of less choice-constrained households can set the constraints for others less able to exercise their housing choices.

As mentioned, it is not possible to make firm conclusions about constrained choices on social frontiers just from movement data. What we can do is consider a range of possibilities and examine what the results tell us about them. Turnover is a particularly useful lens for thinking about this. Low turnover can mean two quite opposite things: relatively well-off, stable communities where sales do not happen often, or poorer communities facing tougher constraints, unable to move (Fotheringham et al., 2004, p. 1670). High turnover levels for 19 to 37 year old native individuals on the native frontier side appears to suggest fewer constraints – they are able to move – but a lack of desire to stay on that frontier. The lack of the same dynamic for older 38 to 65 year old natives could either be due to higher constraints in that group, or because they are happier on the frontier.

An examination of differences in household wealth on frontiers could be informative here: analysis of Rotterdam data suggests that native households on frontiers with high-native proportions tend to be poorer than native households in similar native-mix zones *not* on frontiers (the same is not true for nonwestern households). This is suggestive of constrained options for households on more-native frontiers. Turnover numbers indicate younger native individuals may be leaving households – though perhaps the rest of the household remains behind. Note that the pattern is similar if broken down by ethnicity of household: native households on the frontier, if in higher-nonwestern-proportion zones, are not distinguishable in terms of wealth from others. This could suggest sorting behaviour: those households may have different preferences to those in more native frontier areas. However, while this may suggest constrained choice –

native households less able to move away from frontiers due to financial constraints – it is difficult to separate out wealth differences of the area frontiers are generally located. This suggests a fruitful area for further research, particularly if it was combined with a deeper look at choices on frontiers using survey data.

Neighbourhood allegiance

Turnover is, again, a good place to start for thinking about the neighbourhood allegiance hypothesis, or other ideas about the symbolic and cultural impact of frontiers on mobility. If people were staying on frontiers due to neighbourhood allegiance, turnover would be low – as indeed it is for the majority of results. But the different results for 19 to 37 year old natives on the native frontier side (higher turnover relative to the non-frontier areas of the city) is instructive: if that result is correct, allegiance there would be weak for native households, suggesting – for those households staying – constraints are stronger than any symbolic pull to stay.

So, whilst it is difficult in this data to disentangle symbolic frontier effects from ethnic-density homophily and constrained choice effects, it is clear that, if there are any symbolic effects, they are not symmetric, they vary by age group and are perhaps not as strong as the force of constrained choice. For the younger native age group where a symbolic effect might *a priori* be expected to be strong does not appear to be present.

7. Conclusion

Social frontiers clearly display unique mobility dynamics compared to non-frontier areas, with strong evidence of asymmetries. Our results are consistent with the idea of social frontiers generating "conflicting geographies". More generally, our findings highlight social frontiers as an important area for future study and a potentially enlightening prism for thinking about mobility dynamics in the context of ethnicity and segregation more generally.

We might expect natives on the native side of the frontier – "own-side" frontier dwellers – are likely to have both the motive and the opportunity to move away. They are likely to have larger choice sets than nonwesterners in terms of alternative locations to move to, and may be more likely to view frontiers as a threat (as per the "white flight" narrative of segregation). We would therefore expect native own-side frontier dwellers to be more likely to move away, and this is what we find in our results.

For nonwestern own-side frontier dwellers the opposite seems to be true. This may be because they view frontiers as an opportunity to live close to native households and yet remain within the safety of their own community. So for nonwesterners, own-side frontier zones may be viewed as desirable locations, making residents less likely to move, which again is what we find.

For natives who are "other side" frontier dwellers, it is possible that they tend to be heterophilous (diversity loving) and this is why we find that they are less

likely to move away. However, it is possible that non-westerners on the native side may find it a hostile environment even if they are heterophilous. For example, they may be more likely to feel ostracised by their neighbours if they live on the native side of the frontier. If so, it would explain why we find they are more likely to relocate.

Clearly, this narrative interpretation of our findings is somewhat stylised and further research would be needed to verify these interpretations, perhaps drawing on qualitative and ethnographic methods. Developing causal inference strategies to disentangle the various factors that drive the net outcomes observed in our models would also be a valuable complement to our study. It would also be worth considering other types of social frontier, including "intersectional frontiers" – frontiers that overlap in multiple dimensions such as ethnicity, income, housing tenure and wealth – and comparing how they affect a range of outcomes not just moving behaviour (e.g. wellbeing, educational achievement, household composition).

As yet, we do not know how context mediates the meaning that residents ascribe to frontiers, and so there is a wide field of inquiry to explore with respect to how contextual factors determine the variation of frontier impacts across cities, regions, countries and time periods. We have explored the implications of social frontiers for a particular geography (100m grid squares) but it's possible that the impacts of social frontiers vary for different scales of areal units and for different areas. These variations could, for example, be explored using multilevel models and spatial varying regression models respectively.

The effects of physical boundaries are also an interesting factor that needs further exploration. Noonan (2005) suggests that natural and manmade barriers serve to reduce outgroup exposure and therefore mitigate the negative externalities associated with proximity to outgroups. This in turn might encourage frontiers to form and persist around such barriers. However, the reduced exposure offered by physical barriers may also mitigate frontier impacts (Iyer and Pryce 2022). If so, then it is likely that our results provide a lower bound (i.e. and underestimate) of the impacts of social frontiers. Exploring how frontier impacts vary between stable and fleeting frontiers, and between frontiers with and without physical barriers, would be another useful avenue for future research.

References

Allport, G.W. (1954) The Nature of Prejudice. Addison-Wesley, Chicago.

Bakens, J., Pryce, G. (2019) Homophily horizons and ethnic mover flows. Housing Stud. 34, 925–945.

Binnenlandsbestuur (2020) Rotterdam groeit harder, 2nd March 2020, Accessed 3rd May 2022: www.binnenlandsbestuur.nl/financien/economische-groei-rotterdam-harder-dan-gemiddeld

Clark, W.A.V. (1989) Revealed Preferences and Neighbourhood Transitions in a Multi-Ethnic Setting. Urban Geog. 10, 434–448.

Clark, W.A.V. (1991) Residential Preferences and Neighborhood Racial Segregation: A Test of the Schelling Segregation Model. Demog. 28, 1–19.

Dean, N., Dong, G., Piekut, A., Pryce, G. (2019) Frontiers in Residential Segregation: Understanding Neighbourhood Boundaries and Their Impacts. Tijdschrift voor economische en sociale geografie 110, 271–288.

Dieleman, F.M. (2001) Modelling residential mobility; a review of recent trends in research. J. Housing Built Env. 16, 249–265.

Dennett, A., Stillwell, J. (2008) Population turnover and churn: enhancing understanding of internal migration in Britain through measures of stability. Population Trends; Basingstoke 24–41.

Easton, S., Pryce, G. (2019) Not so welcome here? Modelling the impact of ethnic in-movers on the length of stay of home-owners in micro-neighbourhoods: Urban Studies.

Fotheringham, A.S., Rees, P., Champion, T., Kalogirou, S., Tremayne, A.R. (2004) The Development of a Migration Model for England and Wales: Overview and Modelling Out-Migration. Environ Plan A 36, 1633–1672.

Ibraimovic, T., Hess, S. (2017) Changes in the ethnic composition of neighbourhoods: Analysis of household's response and asymmetric preference structures. Papers Reg. Sci. 96, 759–784.

Iyer, A., Pryce, G. (2022) Theorising the Causal Impacts of Social Frontiers, Nordforsk Working Paper.

Kramer, R. (2017), Defensible Spaces in Philadelphia: Exploring Neighbourhood Boundaries Through Spatial Analysis. RSF: Journal of the Social Sciences, 3, pp.81-101.

Legewie, J., Schaeffer, M. (2016) Contested Boundaries: Explaining Where Ethnoracial Diversity Provokes Neighbourhood Conflict. Am. J.Sociology 122, 125–161.

Maguire, A., French, D., O'Reilly, D. (2016) Residential segregation, dividing walls and mental health. J Epidemiol Community Health 70, 845–854.

Noonan, D.S. (2005), Neighbours, Barriers and Urban Environments: Are Things 'Different on the Other Side of the Tracks'? Urban Studies 42, pp. 1817–1835.

Onderzoek (2022) Bevolking: Rotterdam, Accessed 3rd May 2022: www.onderzoek010.nl/dashboard/onderzoek010/Bevolking/

Piekut, A., Pryce, G. (2022) Life at the Social Frontier: Re-conceptualisation, synthesis, and future directions, NordForsk Working Paper.

Sack, R. D. (1983). Human territoriality: a theory. Annals of the association of American geographers, 73(1), 55-74.

Schelling, T.C. (1978) Micromotives and Macrobehavior. W. W. Norton & Co.

Spielman, S. E. & J. R. Logan. (2013), Using High-Resolution Population Data to Identify Neighbourhoods and Establish Their Boundaries. Annals of the Association of American Geographers, 103, pp. 67–84.

Turner, Frederick Jackson (1893) The Significance of the Frontier in American History. Penguin Books Ltd. 2008 Kindle Edition.

Yinger, J. (1976), Note on the Length of the Black-White Border. Journal of Urban Economics 3, pp. 370–382.

Zorlu, A., Mulder, C.H. (2008) Initial and Subsequent Location Choices of Immigrants to the Netherlands. Regional Studies 42, 245–264.

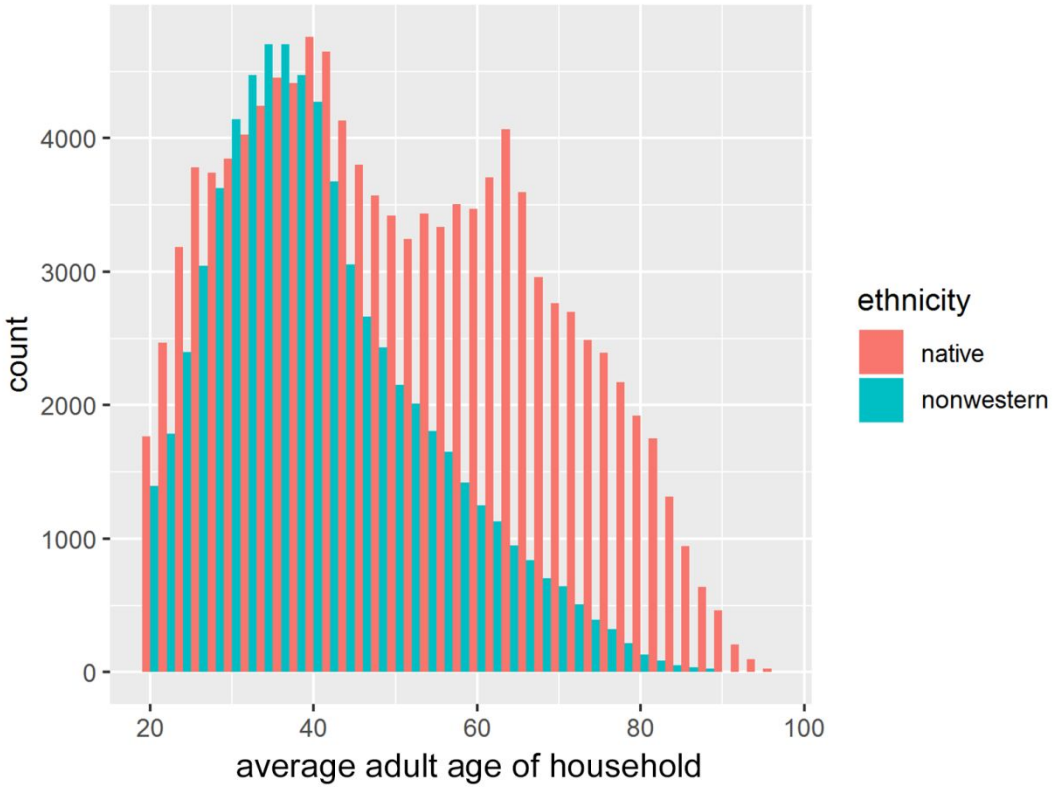


Figure 1: Counts of households in Rotterdam by average adult age (in 2 year bins) and ethnicity (native / nonwestern)

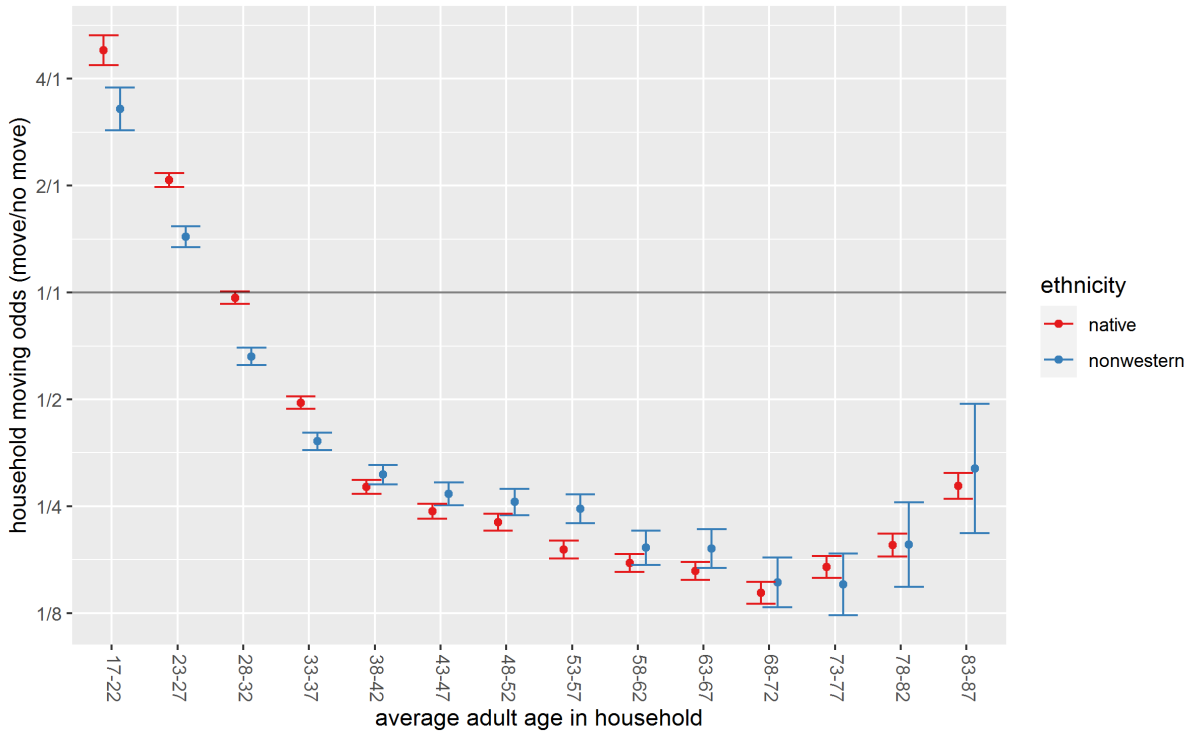


Figure 2: Household moving odds by ethnicity & 5 year age band. Line is even odds (1 to 1) of staying. Above the line is higher odds of moving; below is lower.

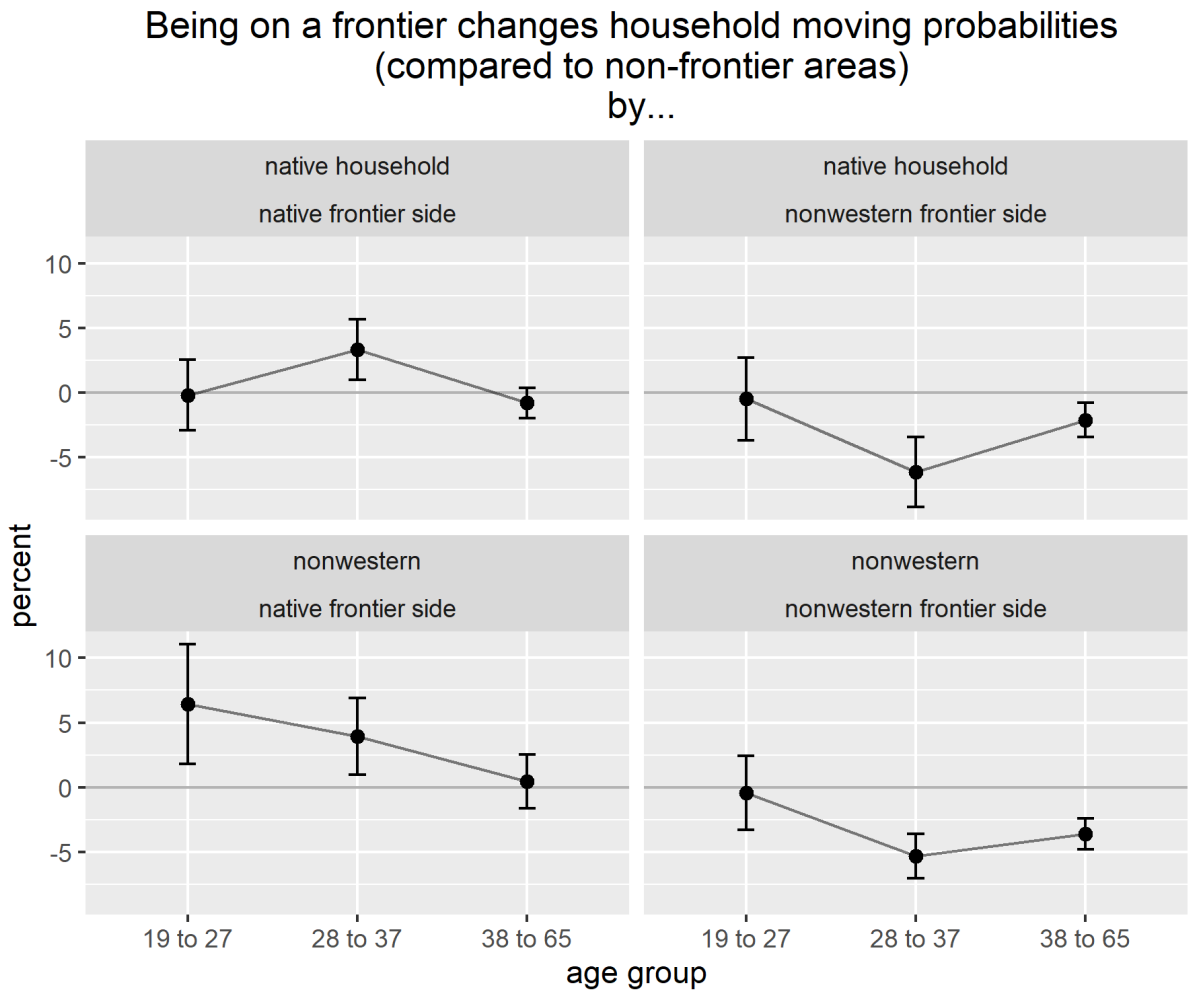


Figure 3: how do social frontiers affect household moving probabilities? Top row shows native households, bottom row nonwestern households. Left-hand column shows those households on the more-native frontier side, right-hand column the more-nonwestern frontier side. Three average-adult-age groups are on the x axis. The y axis shows how much moving probabilities change on those frontier sides compared to non-frontier areas of the city.

The Conflicting Geographies of Social Frontiers

Supplementary Material

Contents

1. Background information on Rotterdam 1

2. Data: Information on the Sample 2

3. Background to empirical estimation of social frontiers 3

4. Sensitivity Analysis 5

Q1/ What is the Impact on Coefficient Estimates of varying the cut-off value for Phi? 5

Q2/ What is the Impact on Coefficient Estimates of removing isolated frontiers? 8

Q3/ What is the Impact on Coefficient Estimates of Introducing Separate Time Periods for Frontier Identification and Frontier Impact Estimation? 10

5. Regression Coefficients (Linear Probability Models) 14

6. Logit Regressions 15

7. Flow models of individual movement behaviour 17

Estimating difference on frontiers compared to non-frontier areas of the city 17

Control Variables 17

8. Migration efficiency 18

9. Turnover 20

1. Background information on Rotterdam

The Netherlands has a relatively recent history of large-scale migration compared to other European countries like the UK, although segregation is not new (Lesger and Van Leeuwen, 2011). The city of Rotterdam is the second largest city of the Netherlands with more than 600,000 inhabitants within the city boundaries, and more than 1 million inhabitants in the larger region. The Netherlands has a relatively recent history of large-scale migration compared to other European countries like the UK, although segregation is not new (Lesger and Van Leeuwen, 2011). The city of Rotterdam is home to one of the largest ports in the world, and this port is a major feature of the geography of the city. The city is divided by the river Nieuwe Maas in a northern and a southern part.

Rotterdam-South is the poorest urban district of the Netherlands, and home to more than 200,000 people. It is the only urban district which has a nationally funded renewal program (National Program Rotterdam South), which aims to lift the level of education and income in this part of the city to the average level of the four largest cities in the country within 20 years. Although the program is very successful, this goal is hard to reach because the average of the four cities is increasing due the relative success of other larger cities in the country.

2. Data: Information on the Sample

As the data only records individuals and households as present at different timepoints and places, if a specific record is no longer present in the data it is impossible to know if that person or household is missing from the data or has left the country. Because of this, only households that moved somewhere within the Netherlands are used to avoid , so that any missing data being is not interpreted as “moved”. We use only households containing a single ethnicity.

Summary statistics for households in the sample:

1. Count of households by ethnicity and whether household moved in the 6 year time period:

	Native	**Nonwestern**
Moved	76535	43783
No move	28411	18803

2. Table 1 in percents per ethnicity:

	Native	**Nonwestern**
Moved	72.9%	70%
No move	27%	30%

3. Count of households living on different sides of frontiers versus living in other parts of Rotterdam, by ethnicity:

	Native	**Nonwestern**
Not on frontier	86543	48604
On “own side” of frontier	10818	10656
On “other side” of frontier	7585	3326

4. Table 3 in percents per ethnicity:

	Native	**Nonwestern**
Not on frontier	82.5%	77.7%
On “own side” of frontier	10.3%	17%
On “other side” of frontier	7.2%	5.3%

3. Background to empirical estimation of social frontiers

We draw on recent developments in the Bayesian areal Wombling literature (see Jacquez et al. 2000[1]; Lu & Carlin 2005[2]; Lee and Mitchell 2013[3]) to estimate social frontiers and to compute inference for our estimates. We not only approximate the locations of frontiers, but we also compute the probability that these estimates are not just the product of random variation in the data. Computing inference is particularly problematic in the presence of spatial autocorrelation which is prevalent in demographic data such those used in the current study. A second issue is that many approaches to boundary detection (e.g. cluster methods) impose "closed" boundaries. It is important, however, in our application to allow for "open" boundaries. This is important because in reality a social frontier may not enclose an entire neighbourhood; it may, for example, characterise a single section of a community's perimeter. Indeed, this is what we find in our data.

Our approach is to apply the Dean et al. (2019) method which addresses both of these issues. We use a Bayesian spatial autoregressive model to capture the uncertainty and overall spatial smoothness of the data. Boundaries are identified through an iterative process which entails making the contiguity matrix endogenous to the estimation process. Places where large step changes in the data occur are deemed to have a low correlation (low similarity) with the neighbouring area, and si the weights matrix cell is adjusted to reflect this. The model is then re-estimated iteratively until estimates are optimised. This is essentially a locally adaptive spatial conditional autoregressive model where estimation of the spatial model enables the updating of the spatial weights matrix using knowledge of the boundaries that have been discovered. To reduce processing time, the spatial autoregressive model is estimated by employing an approximate Bayesian inference approach called an integrated nested Laplace approximation (INLAs, Rue et al. 2009[4]).

We then introduce a second step to ensure that the boundary is not just statistically significant from zero, but also substantively different from zero. This entails setting a threshold of one standard deviation for the estimated probability that two geographical zones have a frontier between them. So if a social frontier identified in the first step has an associated probability below this threshold, we reject it as a social frontier in the second step. On the other hand, if the probability is greater than the threshold, we go ahead and label it as a frontier in the second step. We also test how our regression results for mobility impacts vary if we change the value of this threshold – see Sensitivity Analysis results in the next section. For more information on the Bayesian estimation method see Dean et al. (2019),[5] and Lee and Mitchell (2013)[6].

A map of the proportion non-western in Rotterdam, with social frontiers overlaid in red, is presented below in Fig. S1.

---

[1] Jacquez, G. M., S. Maruca & M. J. Fortin (2000), From Fields to Objectives: A Review of Geographic Boundary Analysis. Journal of Geographical Systems 2, pp. 221–241

[2] Lu, H & B.P. Carlin (2005), Bayesian areal Wombling for Geographical Boundary Analysis. Geographical Analysis, 37, pp. 265–285.

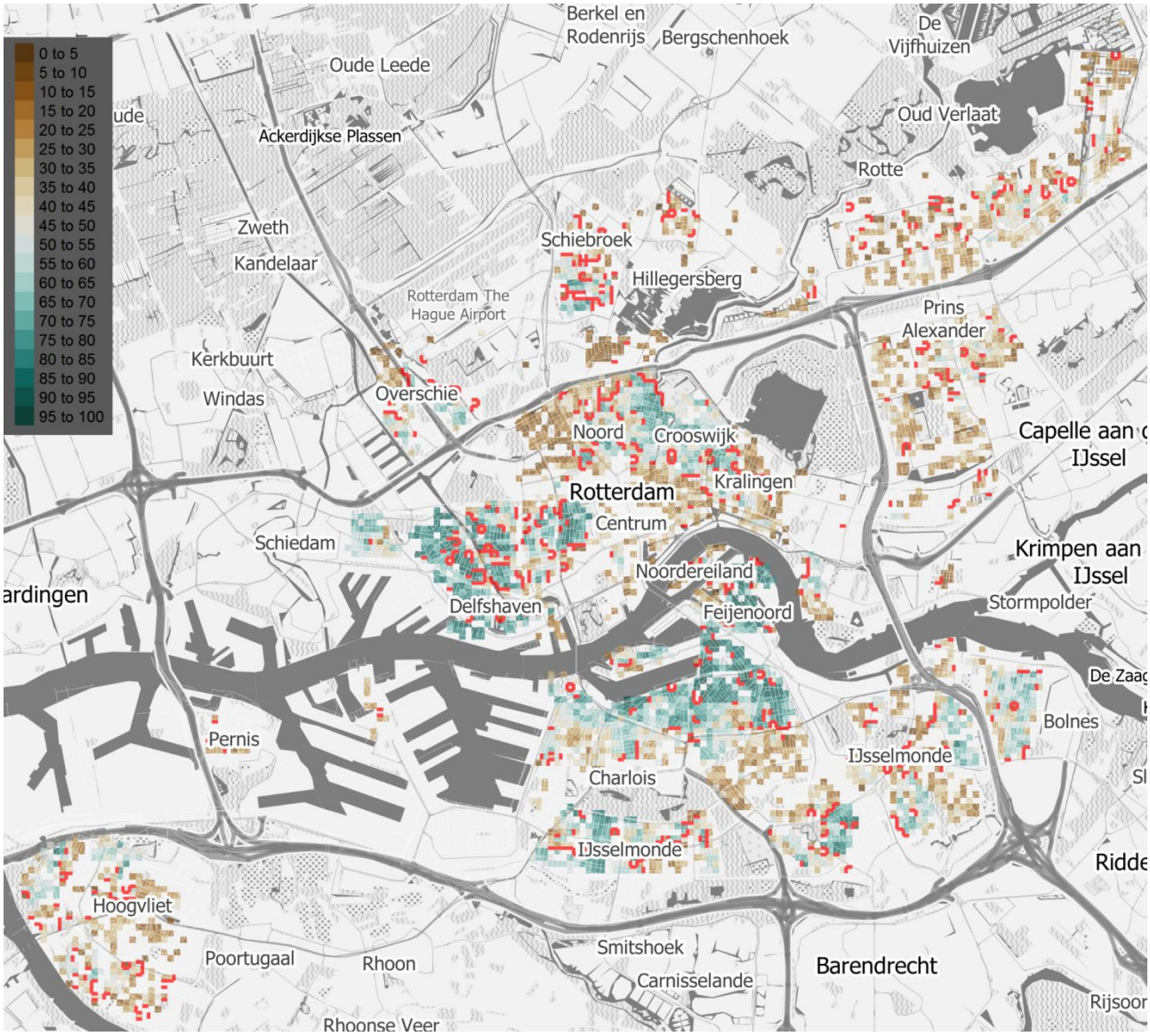
[3] Lee, D. & R. Mitchell. (2013), Locally Adaptive Spatial Smoothing Using Conditional Auto-regressive Models. *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 62, pp. 593–608.

[4] Rue, H., S. Martino & N. Chopin (2009), Approximate Bayesian Inference for Latent Gaussian Models by Using integrated Nested Laplace Approximations (With Discussion). Journal of the Royal Statistical Society, Series B (Statistical Methodology) 71, pp. 319–392.

[5] Dean, N., Dong, G., Piekut, A., & Pryce, G. (2019). Frontiers in residential segregation: understanding neighbourhood boundaries and their impacts. *Tijdschrift voor economische en sociale geografie,* 110(3), 271-288.

[6] Lee, D. & R. Mitchell. (2013), Locally Adaptive Spatial Smoothing Using Conditional Auto-regressive Models. *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 62, pp. 593–608.

Fig. S1: Map of Social Frontiers in Rotterdam with Major Roads & Key Place Names, with percentage nonwestern and frontiers overlaid



The map shows 100m^2 grid squares in Rotterdam coloured by the percentage of nonwestern people, as a proportion of nonwestern plus native people in that zone. Due to secure data disclosure requirements, grid square percentage data is only included for zones where both native **and** nonwestern counts are above 10 people. Areas with, for example, very small counts of nonwestern people do not appear - but these are not in areas where frontiers are found. Frontiers are overlaid in red; these are frontiers used in the main paper's model: stable frontiers from 2012 to 2017, phi = 1. There are no disclosure restrictions on frontiers, so these are all included on the map.

4. Sensitivity Analysis

We consider the impacts on coefficient estimates of three changes:

- *varying the cut-off value for Phi*
- *removing isolated frontiers*
- *separating the time periods for frontier identification and frontier impact estimation*

More detail on these questions is given below. In the write-up we focus on describing the impact on the regression coefficients for the 18-37 age group for whom our main results were most the most significant.

Q1/ What is the Impact on Coefficient Estimates of *varying the cut-off value for Phi*?

First, some background notes on what phi is and why it's important:

- In our main analysis we estimate frontiers between native Dutch households and nonwestern migrant households using a spatial Bayesian model which returns a probability distribution describing the likelihood that two geographical zones have a frontier between them (Dean et al., 2019). These probability values are then normalised and labelled Φ – i.e. the greek letter *phi*.
- We then use a double-hurdle approach to identifying whether a frontier exists between two neighbourhoods.
    - First we select boundaries where the model identifies the spatial discontinuity between between two neighbourhoods to be statistically significant at the 5% level.
    - Second, from these statistically significant discontinuities, we select only those that are substantively different. We do this by choosing only those statistically significant frontiers that also have a value of phi, the probability that the two geographical zones have a frontier between them, that is above a threshold equal to *alpha* (□) standard deviations (σ) from the mean value of phi (ϕ):

$$\text{Border segment is a Social Frontier iff:} \quad \phi \geq (E(\phi) + \square.\sigma)$$

- For comparability, we set *alpha* = 1, which is the threshold used in Dean et al. (2019).

So the question for our sensitivity analysis is the extent to which the Dean et al. (2019) choice of cut-off value for phi (= 1 standard deviation from the mean) affects our main results. Intuitively, we would want to avoid setting the threshold too high as this would only leave a few extreme frontiers. It would also mean that lesser, but potentially important, spatial discontinuities would be bundled in with the "no-frontier" category. Conversely, if we set the threshold too low, even modest spatial discontinuities would be categorised as social frontiers.

To investigate, we re-run our entire range of results for a spectrum of phi thresholds from 0.0 to 4.0 to see how sensitive our regression coefficients are to varying this cut-off value for phi. The results are depicted in Fig.S2 and explained in the bullet points below for the 18-37 age group.

“Own-side” frontier dwellers:

- *Natives on the native side of the frontier (18-37 age group):*
    - In our main results, being in a frontier area had a significant positive effect on the probability of moving for this group.
        - *Impact of varying phi threshold:* Looking at the middle panel of Fig.S2 for the relevant group, we see this remains true for values of the phi cut-point between 0.2 and 1.8 after which the coefficient becomes unstable and insignificant (but typically stays above 0).

- *Nonwesterners on the nonwestern side of the frontier (18-37 age group):*
    - In our main results, being in a frontier area had the opposite effect for “own-side” nonwesterners than for “own-side” natives: i.e. for own-side nonwesterners, proximity to a frontier reduces likelihood of moving.
        - *Impact of varying phi threshold:* We can see from Fig.S2 that this finding holds true for all values of the phi cut-point. Estimates of the frontier coefficient are all negative across the sweep of thresholds for phi. The magnitude of this negative effect increases as the phi threshold rises, but so do the confidence intervals. Nevertheless, the coefficients remain significantly negative for all values of the phi cut-point except for phi = 2.0 even though the confidence intervals widen considerably for phi > 2.2. Overall, this strongly confirms our asymmetry hypothesis that own-side impacts of frontiers differ markedly between natives and nonwesterners.

“Other-side” frontier dwellers:

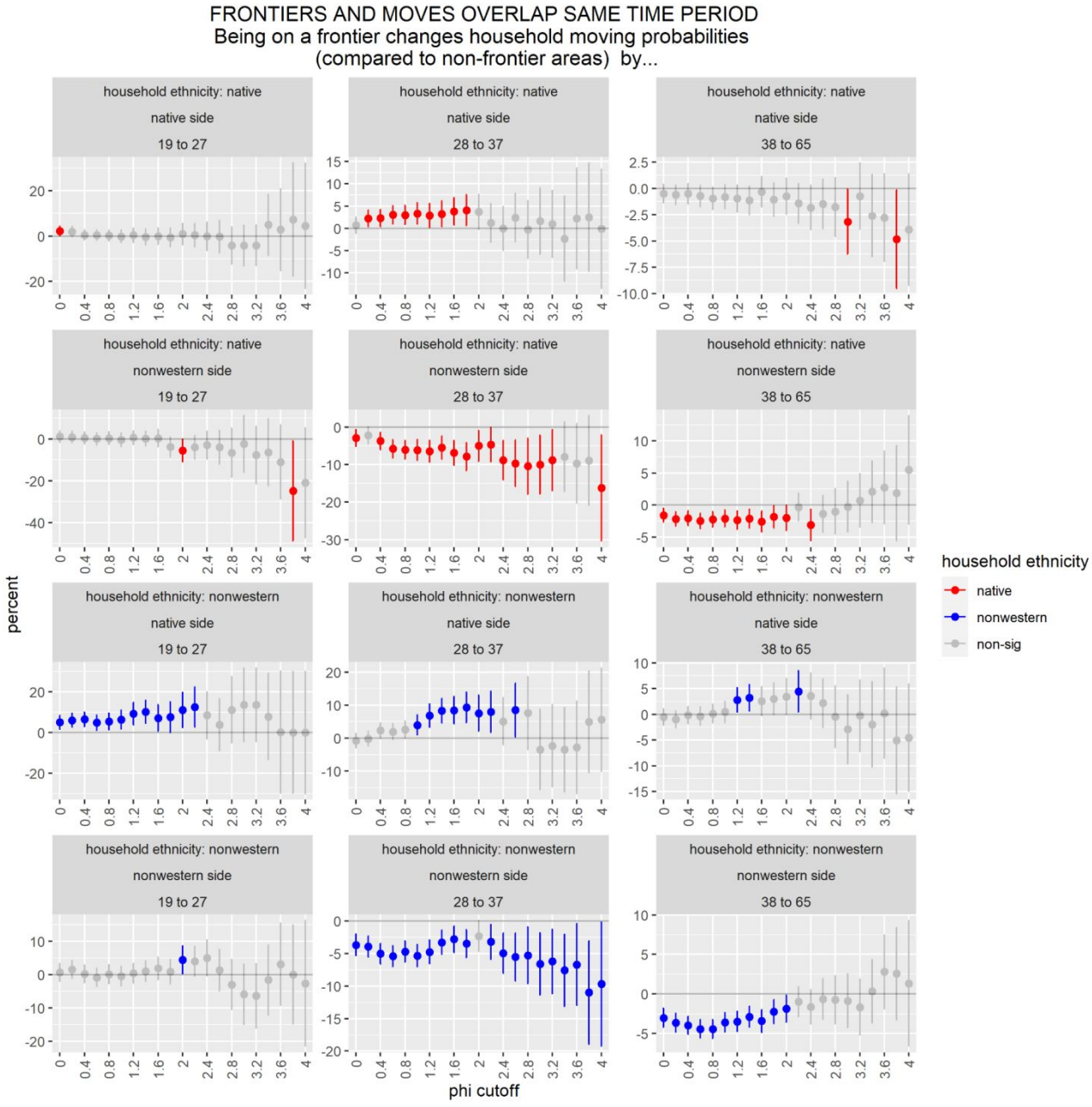
- *Natives on the nonwestern side of the frontier (18-37 age group):*
    - In our main results, the effect of being in a frontier area had a statistically significant negative effect on moving probabilities for “other-side” natives (i.e. natives on the nonwestern side).
        - *Impact of varying phi threshold:* Again, our main finding holds true across all values of the phi cut-point, and significantly so for values of the phi threshold other than 3.4, 3.6, and 3.8.

- *Nonwesterners on the native side of the frontier (18-37 age group):*
    - In our main results, being on the “other side” of the frontier had the opposite effect for nonwesterners compared with natives. The effect of being in a frontier area had a statistically significant *positive* effect on moving probabilities for “other-side” non-westerners.
        - *Impact of varying phi threshold:* This finding of a positive impact on moving probabilities for nonwesterners located on the other-side of a frontier holds true for most values of the phi threshold below 2.8, and significantly so for phi cutpoints between 1.0 and 2.2.

Conclusion on the Impact of Varying the Phi Threshold *(18-37 age group):*

In summary, then, the findings presented in the main article assuming a phi threshold of 1 standard deviation, appear to generally hold true for nearly all values of the phi cutpoint below 2.0 standard deviations. Above that threshold, some of the findings (especially for own-side natives and other-side nonwesterners) generally retain the same sign, but become less significant and less stable.

Fig.S2 Impact on Frontier Coefficient Estimates of varying the cut-off value for Phi



Q2/ What is the Impact on Coefficient Estimates of *removing isolated frontiers?*

Our main results produce a number of 'small' frontiers (see Fig. S1) including frontiers that completely surround 1 individual 100m grid cell. These isolates may simply represent noise. It is also possible that they have a less potent effect on moving behaviour compared to longer frontiers.

So the question for our sensitivity analysis is whether removing these single cell frontiers affects our main results? Removing these isolates will reduce the effective sample size of frontier zones which may increase standard errors and confidence intervals, and reduce the statistical significance of our results. It may also truncate our sample in a potentially non-random way. To investigate, we re-ran our entire range of results with the single cell frontiers removed for a spectrum of phi thresholds. The results are visualised in Fig.S3 and explained in the bullet points below for the 18-37 age group.

"Own-side" frontier dwellers:

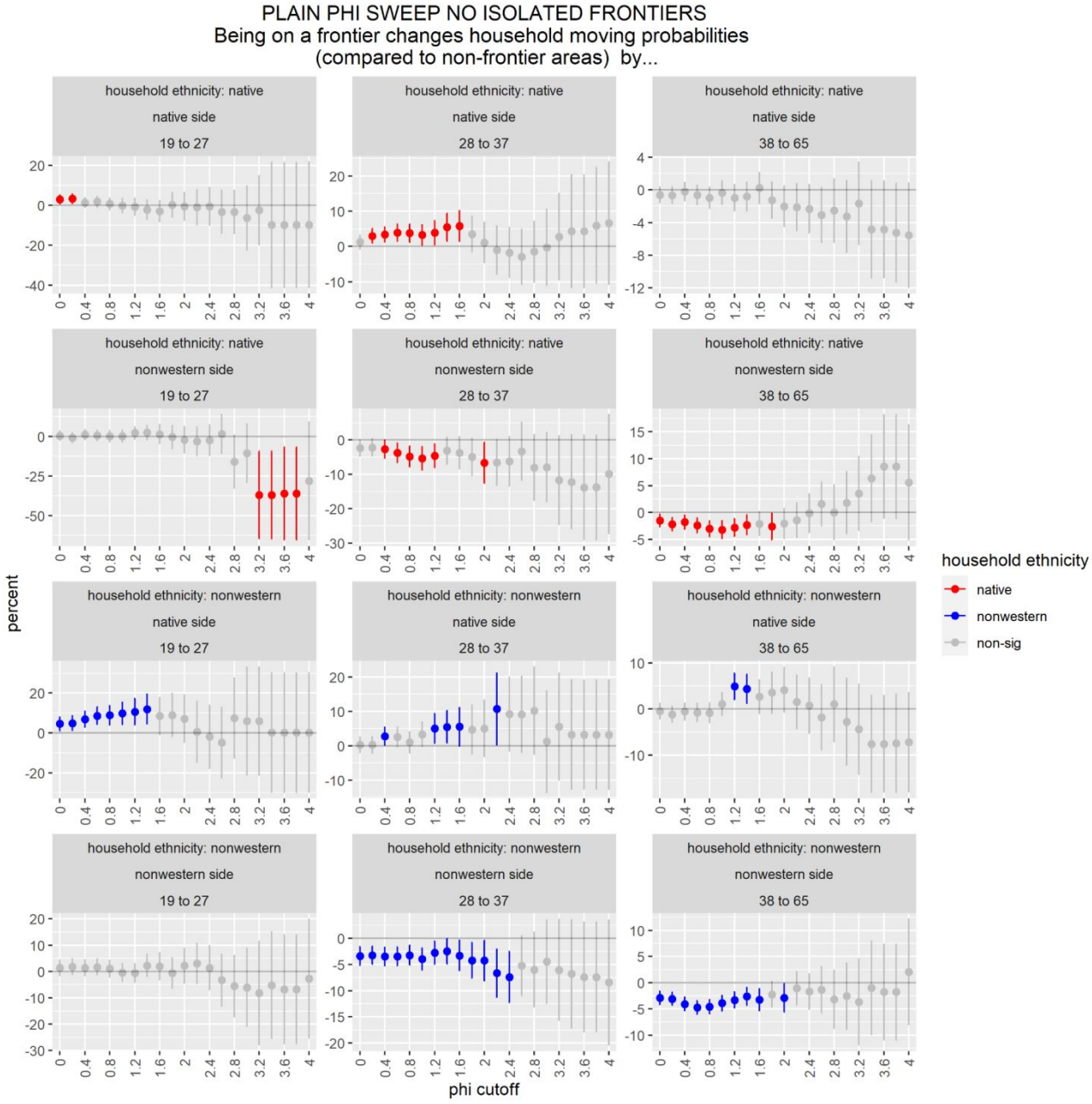
- *Natives on the native side of the frontier (18-37 age group):* when we remove the isolated frontiers (Fig.S3) and compare the results for own-side natives with those from Fig.S2 (isolates included), the frontier coefficient estimates look very similar, at least for phi cut-point values up to 1.6 standard deviations. Beyond that, the estimates become insignificant, dipping below zero (though not significantly so) between the range 2.0 to 3.2 standard deviations.
- *Nonwesterners on the nonwestern side of the frontier (18-37 age group):* The pattern of results for own-side nonwesterners for different phi cut-point values remains pretty much the same whether or not isolates are removed (compare Fig.S2 and Fig.S3). All coefficient estimates for the frontier effect remain below zero, as when isolates are included. The main difference seems to be a reduction in statistical significance (coefficients become insignificant for phi thresholds > 2.4 standard deviations). This may be due to the reduced sample size when we remove isolates.

"Other-side" frontier dwellers:

- *Natives on the nonwestern side of the frontier (18-37 age group):* Again, we find that the pattern of results remains very similar (all coefficient estimates are below zero) but the confidence intervals widen when we remove isolates. Most of the frontier coefficient estimates become statistically insignificant for phi thresholds > 1.2 standard deviations.
- *Nonwesterners on the native side of the frontier (18-37 age group):* Two main effects here. First, we again see a rise in the magnitude of the coefficient but also widening confidence intervals for higher threshold values for phi. Second, the coefficient estimates remain positive (not significantly so) even for higher values of phi. So removing isolated frontiers does not change the main findings, but does make the results less significant especially for phi thresholds above 2.4 standard deviations.

Summary Impact of Removing Isolates *(18-37 age group):*
In summary, then, when we remove isolates, our main findings hold true in terms of the overall pattern of estimates. However, as anticipated, by reducing the sample size for frontier zones, the removal of isolated frontiers has the effect of widening confidence intervals, which makes the estimates less statistically significant, especially for higher values of the phi cut-point.

Fig.S3 Impact on Frontier Coefficient Estimates of Removing Isolated Frontiers



Q3/ What is the Impact on Coefficient Estimates of Introducing Separate Time Periods for Frontier Identification and Frontier Impact Estimation?

The main purpose of our paper is not to provide causal estimation of the impact of social frontiers. Rather, our aim is to describe patterns in moving behaviours and how these vary across different groups for each side of the frontier.

Nevertheless, it would be useful to consider how the results are affected by separating the time period for estimating the location of social frontiers from the period when we estimate their impact.

In the results presented in the main manuscript, we identify social frontiers that are stable for the 6 year period 2012 to 2017. We then use this same time period to estimate social frontier impacts on moving behaviour.

In the sensitivity analysis below, we re-run the phi threshold for estimates of frontier coefficients over the 3 year period 2015-2017 inclusive, based on estimates of stable frontiers for the period 2012-14. Intuitively, we would expect a noticeable widening of confidence intervals due the large reduction in the sample size of movers (i.e. we are now looking at movers over a 3 year period rather than a 6 year one). If we also remove isolated frontiers (Fig.S5) we would expect the widening of confidence intervals to be even more pronounced. The question is whether the magnitudes and signs of the coefficients are markedly different. The results are presented in Fig.S4 (isolated frontiers included) and Fig.5 (isolated frontiers removed) which we now compare with those of Fig.S2.

“Own-side” frontier dwellers:

- *Natives on the native side of the frontier (18-37 age group):* when we estimate household moves in a separate time period to that used for identifying social frontiers, we find very little difference in the pattern of coefficient estimates for different values of the phi threshold (compare the graphs for own-side natives in Fig.S4 and Fig.S5 with Fig.S2). The coefficients remain positive for all values of the phi threshold below 3.2 standard deviations from the mean, significantly so for phi thresholds below 2.0 except for the threshold of 1.4 standard deviations.
- *Nonwesterners on the nonwestern side of the frontier (18-37 age group):* The impact of separating the time period for estimating frontier location and frontier impact is more pronounced for own-side nonwesterners. The frontier coefficient becomes more stable across values of the phi threshold, remaining close to -2.0 throughout the sweep of phi cut-points (Fig.S4 and Fig.S5). However, the separation also reduces statistical significance, presumably because of the reduced sample of movers as a result of the shorter time period for mover estimation. As a result, frontier coefficients that were previously statistically significant, now become insignificant for a number of cut point values (1.4, 1.6, 2.0, 2.2, 2.4 in Fig.S4; even more so in Fig.S5). Nevertheless, the conclusion remains broadly the same albeit with more uncertainty: frontier zones seem to have a negative impact on the moving propensities of own-side nonwesterners, which contrasts with the positive impact for own-side natives.

"Other-side" frontier dwellers:

- *Natives on the nonwestern side of the frontier (18-37 age group):* The frontier coefficient becomes more unstable and less significant when we separate the time periods for estimating frontier location and impact. However, most of the coefficients remain negative even if they are no longer significant.
- *Nonwesterners on the native side of the frontier (18-37 age group):* The frontier coefficient remains positive for most of the range of cut points plotted in Fig.S4, but with a stronger pattern of increasing magnitude beyond the phi threshold of 3.2 standard deviations from the mean.

Summary Impact of Separating Time Periods for Estimation: ***(18-37 age group):***
Separating estimation periods makes some of our results less statistically significant, probably because of the large reduction in the number of moves available for including in the regression due to losing 3 years of data. Nevertheless, the main findings remain broadly the same – the sign of coefficients for the four groups (own-side natives, own-side nonwesterners, other-side natives and other-side nonwesterners) remain largely the same as those in the regressions reported in the manuscript (albeit with increased uncertainty), supporting the hypothesis of asymmetric effects across the frontier.

Fig.S4 Impact on Frontier Coefficient Estimates of Separate Time Periods

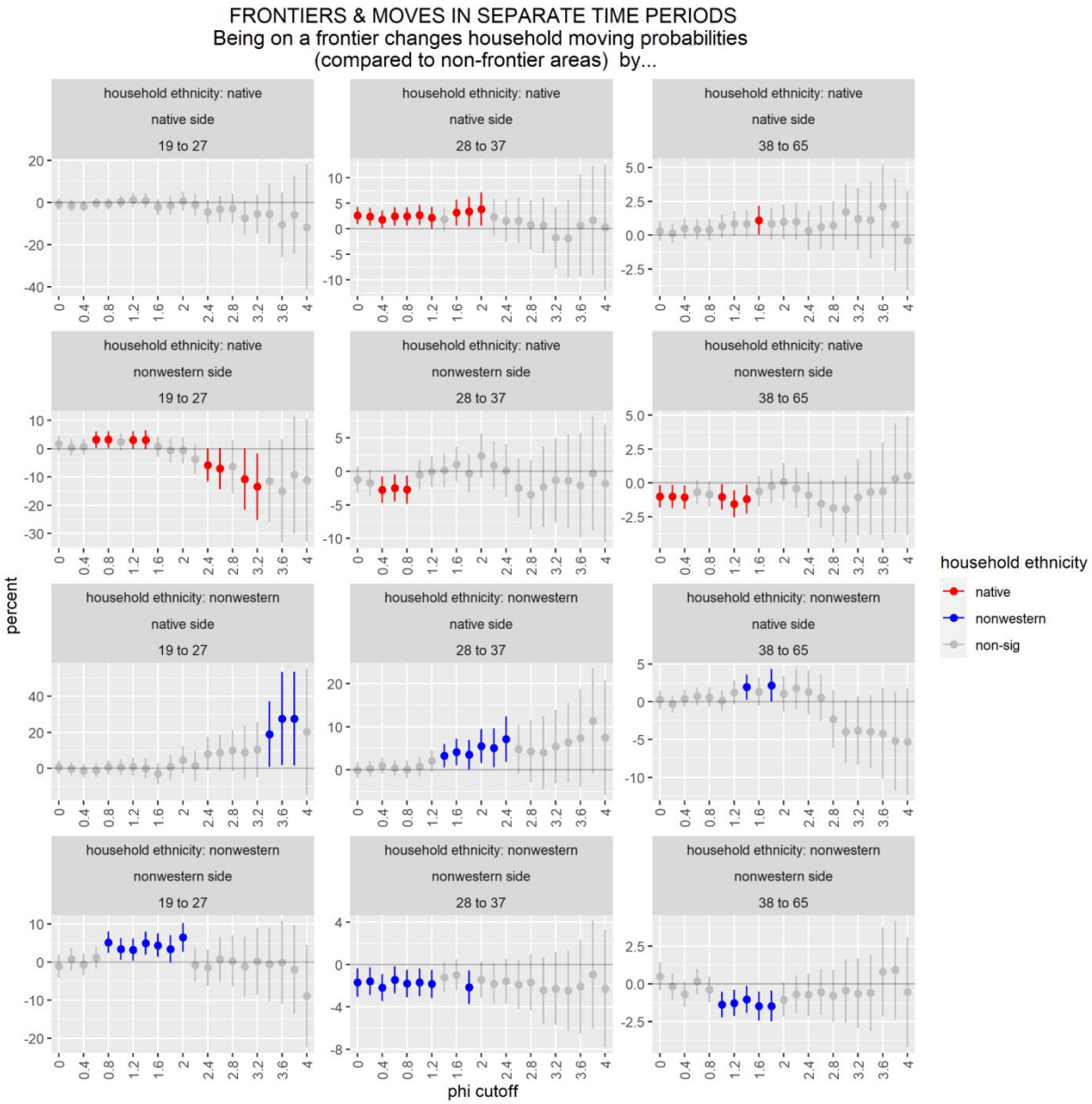
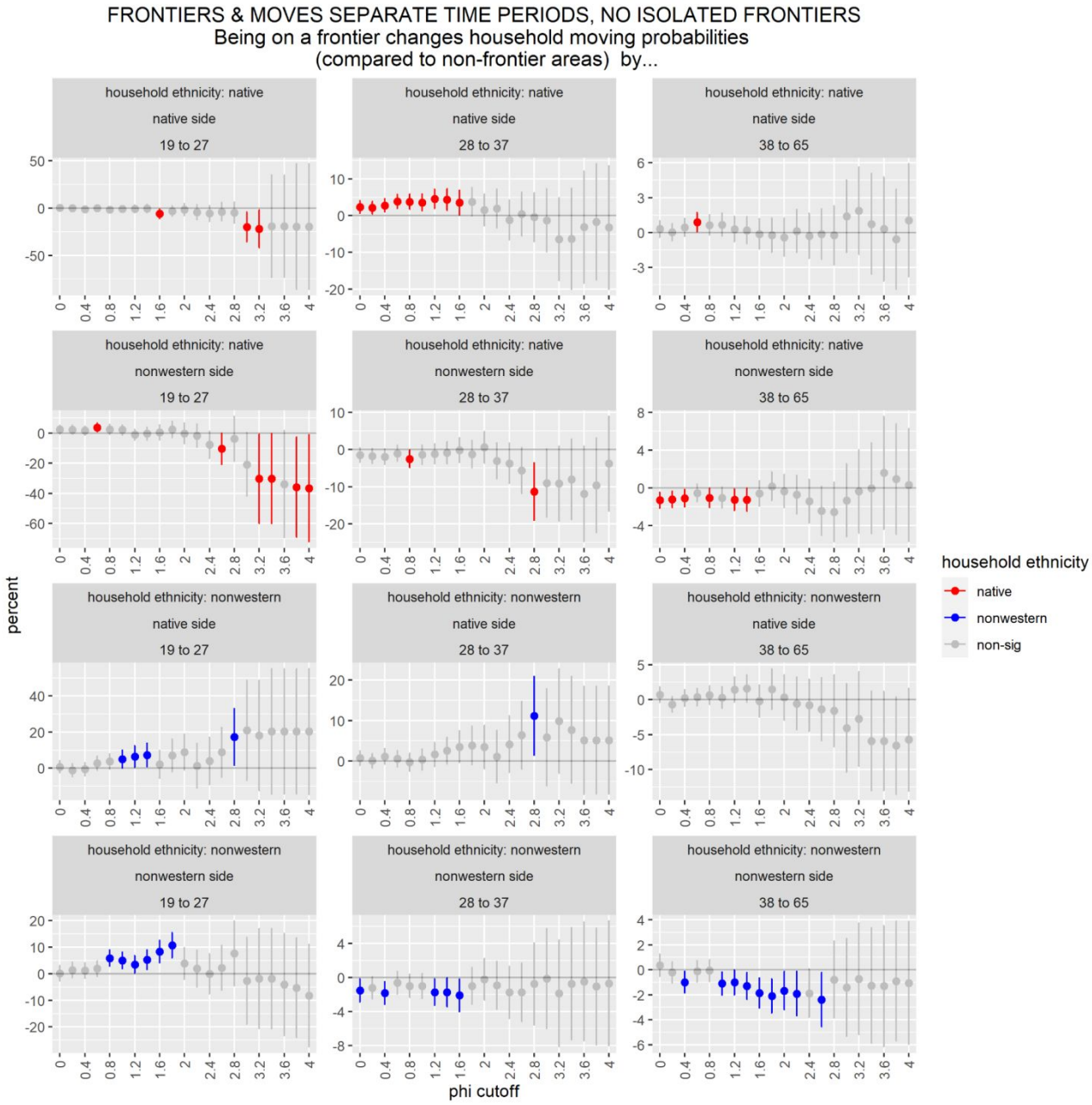


Fig.S5 Impact on Frontier Coefficient Estimates of Separate Time Periods with Isolated Frontiers Removed



5. Regression Coefficients (Linear Probability Models)

Table S1 below provides the full set of coefficients for each of the six main regressions in the paper (see Fig. 3 in the paper) but estimated using logistic regression rather than linear probability models. All coefficients are log odds. Each separate regression is in columns 1 to 6; these are: (1) native, 19 to 27; (2) nonnative-nonwestern, 19 to 27; (3) native, 28 to 37; (4) nonnative-nonwestern, 28 to 37; (5) native, 38 to 65; (6) nonnative-nonwestern, 38 to 65. The first two rows are categorical contrasts for the named ethnicity to the base category of households not on frontiers.

Table S1 Coefficients for the Six Main Regressions, probabilities [0,1]

	Dependent variable:					
	moved					
	(Native 19-27)	(Nonwestern 19-27)	(Native 28-37)	(Nonwestern 28-37)	(native 38-65)	(Nonwestern 38-65)
Frontier: Native side (cf. non-frontier)	−0.002	0.064***	0.033***	0.039***	−0.008	0.005
	(0.014)	(0.024)	(0.012)	(0.015)	(0.006)	(0.010)
Frontier: Nonwestern side (cf. non-frontier)	−0.005	−0.004	−0.061***	−0.053***	−0.021***	−0.036***
	(0.016)	(0.015)	(0.014)	(0.009)	(0.007)	(0.006)
homeownershipRented	0.373***	0.355***	0.245***	0.198***	0.076***	0.085***
	(0.011)	(0.018)	(0.008)	(0.008)	(0.004)	(0.006)
haschildren12to18	−0.052	−0.246***	−0.074***	−0.089***	−0.012*	−0.017**
	(0.106)	(0.052)	(0.015)	(0.009)	(0.006)	(0.007)
standardisedincome_1000s	−0.002***	−0.003***	0.006***	−0.001	0.001***	−0.003***
	(0.0005)	(0.001)	(0.0003)	(0.0004)	(0.0001)	(0.0003)
hse_structureCouple without children	0.067***	−0.035**	0.092***	0.064***	−0.034***	−0.059***
	(0.011)	(0.017)	(0.010)	(0.013)	(0.004)	(0.008)
hse_structureCouple with children	0.011	−0.078***	−0.075***	−0.116***	−0.025***	−0.069***
	(0.024)	(0.025)	(0.009)	(0.009)	(0.005)	(0.007)
hse_structureSingle parent family	−0.220***	−0.194***	−0.169***	−0.196***	0.005	−0.070***
	(0.029)	(0.019)	(0.012)	(0.009)	(0.007)	(0.007)
hse_structureOther multiperson household	0.046*	−0.088***	−0.062	0.021	0.020	0.111***
	(0.026)	(0.020)	(0.043)	(0.020)	(0.022)	(0.023)
prop_owneth10perc_incs	−0.002	−0.003	−0.016***	−0.009***	−0.008***	−0.007***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)
Constant	0.451***	0.407***	0.262***	0.364***	0.177***	0.276***
	(0.019)	(0.025)	(0.015)	(0.015)	(0.008)	(0.011)
Observations	10,688	8,053	19,502	20,200	50,907	30,702
R^2	0.128	0.095	0.098	0.087	0.016	0.028
Adjusted R^2	0.128	0.094	0.098	0.087	0.016	0.028
Residual Std. Error	0.422 (df = 10677)	0.459 (df = 8042)	0.466 (df = 19491)	0.450 (df = 20189)	0.380 (df = 50896)	0.402 (df = 30691)
F Statistic	157.298*** (df = 10; 10677)	84.615*** (df = 10; 8042)	212.583*** (df = 10; 19491)	192.353*** (df = 10; 20189)	81.444*** (df = 10; 50896)	88.849*** (df = 10; 30691)
Note:						*p<0.1; **p<0.05; ***p<0.01

6. Logit Regressions

A number of widely cited papers (Allison 1999; Williams, 2009; and Mood 2010) have argued that "regression models for binary responses are problematic if we want to compare estimated coefficients from models for different groups or with different explanatory variables" (Kuha et al. 2020[7], p.498). As a result, we presented the results of linear probability models (OLS) in the main paper and in the sensitivity analysis above.

For sake of completeness we present the logit results in **Table S2** below which lists the full log odds for each of the six main regressions in the paper (see Fig. 3 in the paper) estimated using logistic regression rather than linear probability models. All coefficients are log odds. Each separate regression is in columns 1 to 6; these are: (1) native, 19 to 27; (2) nonnative-nonwestern, 19 to 27; (3) native, 28 to 37; (4) nonnative-nonwestern, 28 to 37; (5) native, 38 to 65; (6) nonnative-nonwestern, 38 to 65. The first two rows are categorical contrasts for the named ethnicity to the base category of households not on frontiers.

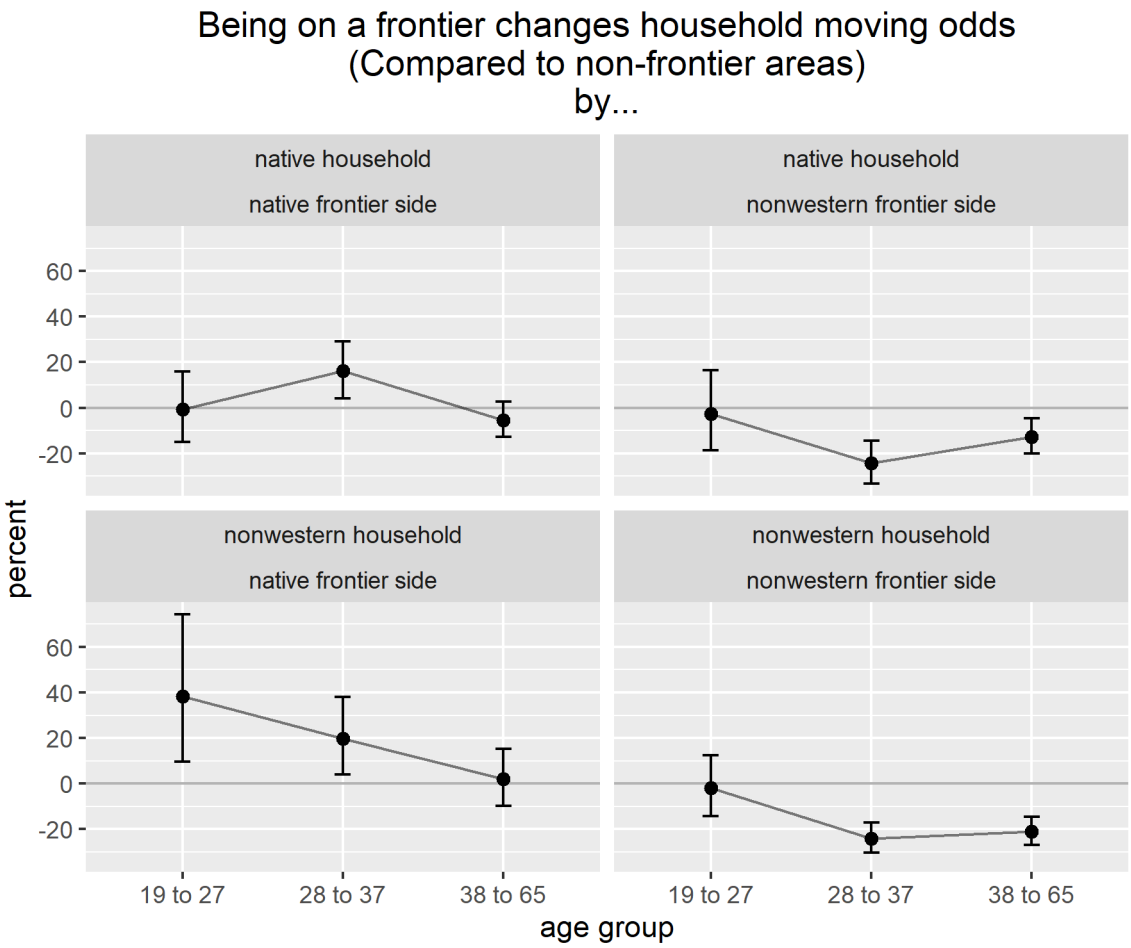
In the event, the pattern of results from the logit regressions are essentially very similar to those from the linear probability models (compare **Fig.3** in the paper with **Fig.S6** below). This confirms Kuha et al's (2020, p.498) conclusion that concerns about using logit models to compare groups are "usually misplaced".

Table S2 Coefficients for the Six Main Regressions, odds model

	Dependent variable:					
	moved					
	(Native 19-27)	(Nonwestern 19-27)	(Native 28-37)	(Nonwestern 28-37)	(native 38-65)	(Nonwestern 38-65)
Frontier: Native side (cf. non-frontier)	−0.008	0.325***	0.149***	0.180**	−0.056	0.020
	(0.079)	(0.118)	(0.055)	(0.072)	(0.042)	(0.062)
Frontier: Nonwestern side (cf. non-frontier)	−0.028	−0.019	−0.280***	−0.274***	−0.136***	−0.235***
	(0.092)	(0.069)	(0.063)	(0.045)	(0.045)	(0.040)
homeownershipRented	1.685***	1.510***	1.091***	1.002***	0.533***	0.600***
	(0.060)	(0.085)	(0.036)	(0.042)	(0.028)	(0.044)
haschildren12to18	−0.296	−1.099***	−0.405***	−0.495***	−0.096**	−0.119***
	(0.601)	(0.256)	(0.080)	(0.047)	(0.047)	(0.045)
standardisedincome_1000s	−0.009***	−0.012***	0.027***	−0.002	0.008***	−0.017***
	(0.003)	(0.004)	(0.002)	(0.002)	(0.001)	(0.002)
hse_structureCouple without children	0.373***	−0.158**	0.395***	0.305***	−0.242***	−0.337***
	(0.064)	(0.078)	(0.045)	(0.060)	(0.032)	(0.051)
hse_structureCouple with children	0.076	−0.354***	−0.342***	−0.514***	−0.185***	−0.415***
	(0.124)	(0.115)	(0.043)	(0.043)	(0.038)	(0.045)
hse_structureSingle parent family	−1.058***	−0.846***	−0.778***	−0.894***	0.040	−0.398***
	(0.142)	(0.084)	(0.057)	(0.044)	(0.049)	(0.046)
hse_structureOther multiperson household	0.264*	−0.400***	−0.282	0.114	0.124	0.619***
	(0.150)	(0.092)	(0.197)	(0.093)	(0.139)	(0.122)
prop_owneth10perc_incs	−0.012	−0.016	−0.073***	−0.044***	−0.055***	−0.040***
	(0.013)	(0.012)	(0.009)	(0.008)	(0.006)	(0.007)
Constant	−0.156	−0.363***	−1.045***	−0.685***	−1.595***	−1.019***
	(0.104)	(0.120)	(0.069)	(0.073)	(0.052)	(0.071)
Observations	10,688	8,053	19,502	20,200	50,907	30,702
Log Likelihood	-5,754.847	-4,914.158	-12,155.360	-11,924.750	-23,473.170	-15,358.250
Akaike Inf. Crit.	11,531.690	9,850.316	24,332.720	23,871.500	46,968.350	30,738.510
Note:						*p<0.1; **p<0.05; ***p<0.01

[7] Kuha, J., & Mills, C. (2020). On Group Comparisons With Logistic Regression Models. Sociological Methods & Research, 49(2), 498–525.

Fig.S6 How do Social Frontiers Affect Household Moving Odds?



Note: Top row shows native households, bottom row nonwestern households. Left-hand column shows those households on the more-native frontier side, right-hand column the more-nonwestern frontier side. Three average-adult-age groups are on the x axis. The y axis shows how much moving odds change on those frontier sides compared to non-frontier areas of the city.

7. Flow models of individual movement behaviour

Estimating difference on frontiers compared to non-frontier areas of the city

We use two complementary flow measures: *migration efficiency* and *turnover* (see e.g. Dennett and Stillwell, 2008). Migration efficiency measures the *polarity of flows* coming into and out of a zone. Bound between -1 and 1, values of -1 indicate all flows were out of the zone, none came in; +1 indicates all flows were into the zone, none left. Zero indicates a balance of inward and outward flows. Migration efficiency provides a measure of polarity but not scale; so we use *turnover* for the latter. The turnover measure is just the sum of inward and outward flows divided by the total population of the zone (actually, an average of the total population at each of the two time points used to acquire the movement data, so that it is consistent). These are used in a standard population flow regression framework (see e.g. Dennett and Stillwell, 2008) to estimate the impacts of social frontiers on population dynamics.

We use individuals rather than households to measure flow, for two reasons. Firstly, it is the raw number of people that is important when considering the impact of flows overall. Second, it maximises observations relative to the count of households; this is needed to provide enough observations for statistically robust results. Relatedly, we add a new category here: all people, including non-native western. This again maximises sample size, giving more significant results than we could otherwise get.

Both of these measures provide a single value for each hundred metre grid square - each grid square is used as an observation in an OLS regression to model flow behaviour on frontiers. Controls are mean-centred so that the intercept gives a point estimate for the flow measures in non-frontier parts of Rotterdam, holding the controls constant. A dummy for whether the zone sits on a social frontier (and which side of the frontier it is) then provides an estimate of the frontier effect on the flow measures. The flow measure regressions are broken further into two groups: nonwestern frontiers are compared to the more-nonwestern half of the city (split by the median); native frontiers are compared to the more-native half.

The two sets of OLS flow and turnover regressions are:

- Dependent variables: one each of migration efficiency and turnover per grid square.
- Explanatory variables:
  - **Social Frontier:** whether the observation grid square is on a social frontier. Also, we identify which side of the frontier the resident is on, the "native side" or "nonwestern side" and compare to the native or nonwestern half of the city.
  - **Age band (average adult age in household):** 19 to 27; 28 to 37; 38 to 65. We also include an "all ages" category to get more statistical power.

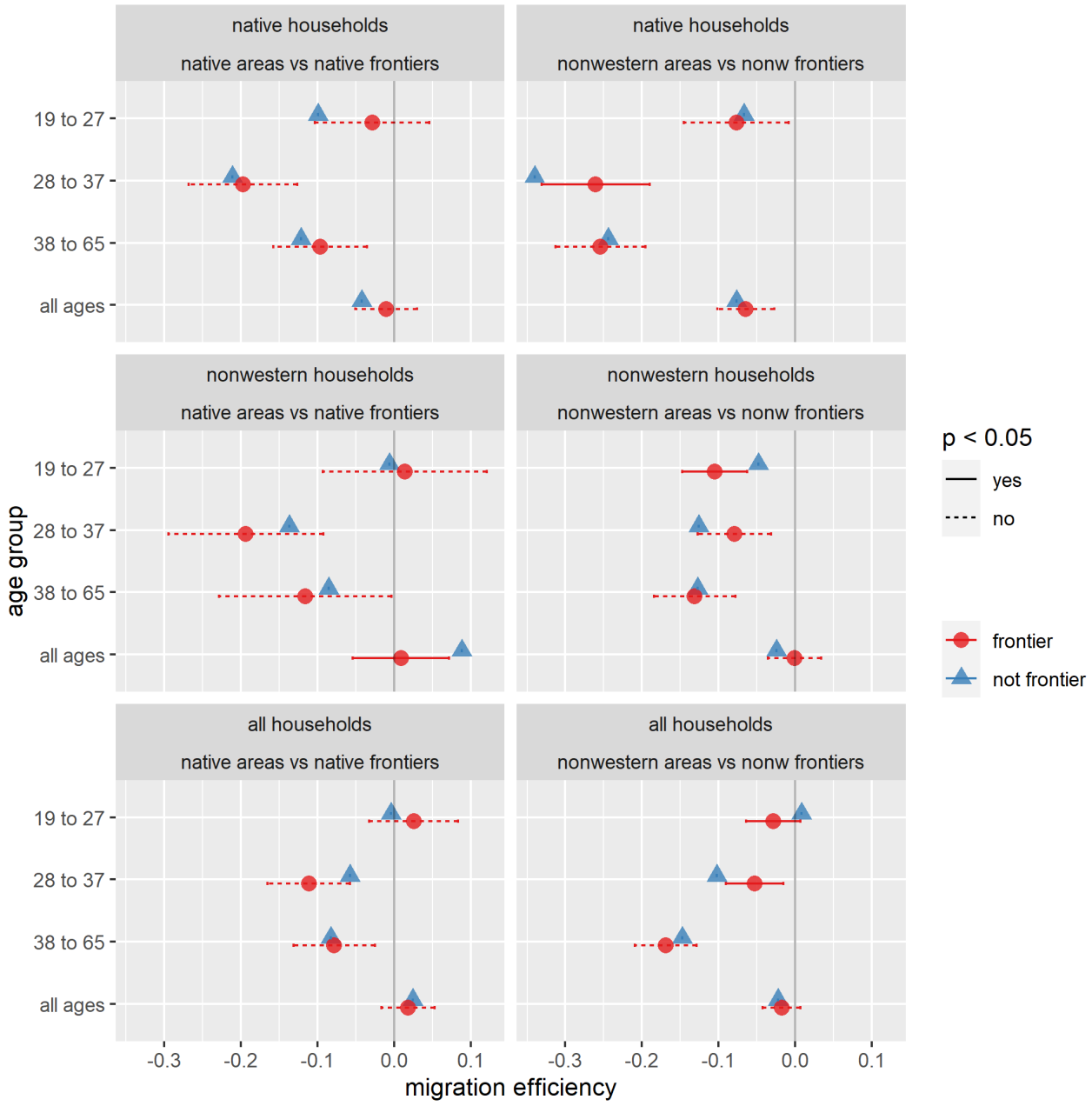
Control Variables

Other variables not part of the hypothesis investigation are as follows:

- **Income**: average standardised income in each grid square
- **Ownership**: proportion of owner-occupied housing in each grid square.

8. Migration efficiency

Fig. S7: Migration efficiency on Frontiers Compared to Non-frontier Areas of the City, Broken down by age group on the y axis (including a category for all ages combined).



Note: Blue triangles give the point estimate for non-frontier areas; red dots with error bars for frontiers – the result is significant ($p > 0.05$) if the triangle is beyond the error bar. Top row: native households. Middle row: nonwestern households. Bottom row: all households. Left-hand column: native frontiers vs more-native non-frontier city areas. Right-hand column: nonwestern frontiers vs more-nonwestern non-frontier city areas

Fig.S7 summarises results for the migration efficiency regressions. The left-hand column of results compares more-native frontiers with more-native non-frontier areas of the city; the right hand column does the same for nonwestern areas and frontiers. Blue triangles are point estimates from the model's intercepts that give the estimated migration efficiency in non-frontier areas of the city; red bars give the result on

frontiers (95% confidence intervals). Results are significant ($p < 0.05$) where the bars do not cross the point estimate: frontier migration efficiency is significantly different from migration efficiency in non-frontier areas of the city. (Non-significant results are indicated by dashed lines). Each panel not only breaks the results down by age group, but also by the pairing of native areas of the city with more-native frontiers and nonwestern areas with non-western frontiers.

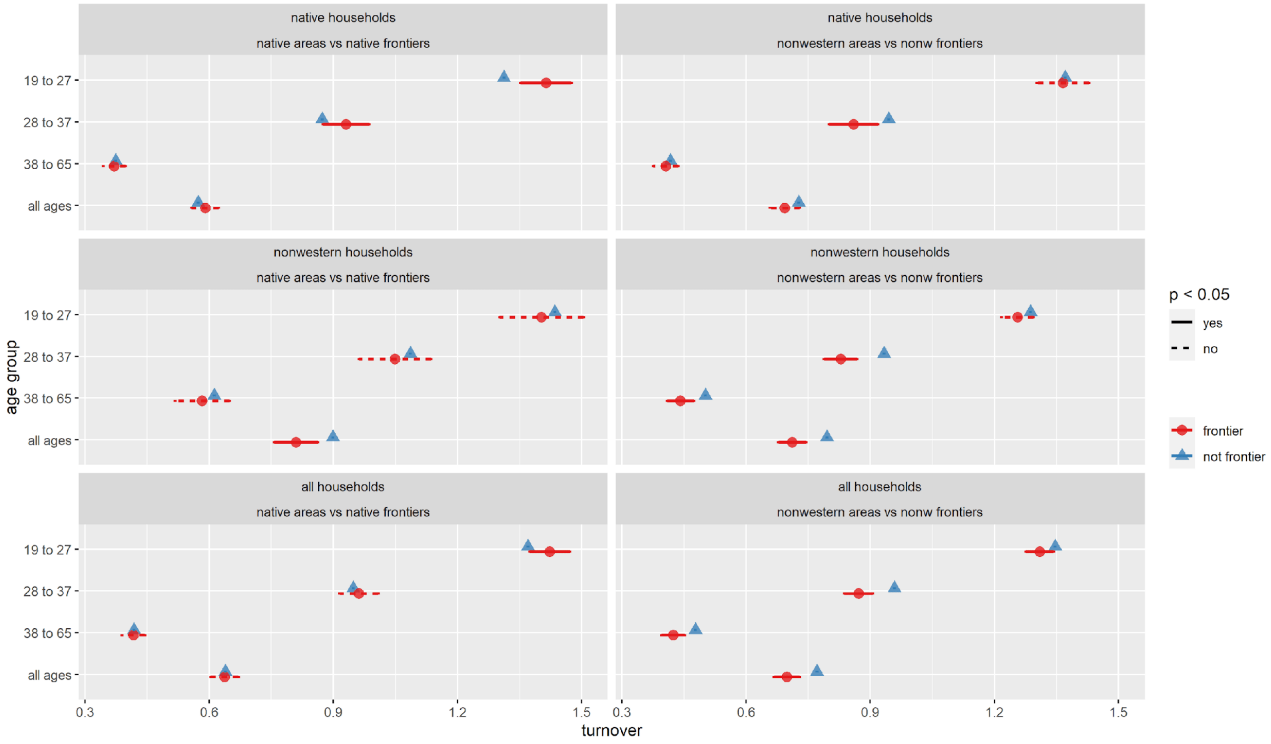
Comparing nonwestern city areas to nonwestern frontiers first, results are opposite for 19 to 27 year old nonwesterners versus 28-37 year old natives. For 19 to 27 year old nonwesterners on the nonwestern frontier side, more leave than come in, relative to the rest of the city (migration efficiency is lower on frontiers). For 28 to 37 year olds, it is opposite: natives on the nonwestern frontier side, more come in than leave, relatively. That pattern is confirmed by the significant result for 'all households', suggesting that – while there are non-significant results – there are different frontier effects on migration efficiency that depend more on age-group than ethnicity.

On the native side of the frontier, there is only one significant result, but an informative one. When modelling all age groups together, nonwesterners on the native frontier side are leaving more than arriving, relative to the rest of the city. The lack of significant results for natives on the native side suggests their number balance may be little different from the rest of the city.

These results are discussed in more detail below, to compare to turnover.

9. Turnover

Fig.S8: turnover on frontiers compared to non-frontier areas of the city, broken down by age group on the y axis (including a category for all ages combined).



Note: Blue triangles give the point estimate for non-frontier areas; red dots with error bars for frontiers – the result is significant ($p > 0.05$) if the triangle is beyond the error bar. Top row: native households. Middle row: nonwestern households. Bottom row: all households. Left-hand column: native frontiers vs more-native non-frontier city areas. Right-hand column: nonwestern frontiers vs more-nonwestern non-frontier city areas

Turnover models (**Fig.S8**) are constructed in the same way as the previous migration efficiency results. The majority of significant findings show similar outcomes: turnover on frontiers is *lower* than in other areas of the city; frontiers are more stable in terms of throughput of people. This is true for: both ethnicities aged 28-37 on the nonwestern frontier side; 38-65 year old nonwesterners on the nonwestern frontier side; when all ages are combined, the result is the same for nonwesterners on both sides of the frontier.

There are a couple of exceptions to this pattern of frontiers being more stable: natives on the native side of the frontier, in the two younger age groups (19-27, 28-37) have significantly *higher* turnover on frontiers compared to other areas of the city. Note in particular the contrast to native 28-37 year olds on the nonwestern side, where turnover is relatively lower; and also to nonwesterners on the native frontier side, when all nonwestern age groups are combined: turnover is relatively lower there also.

Another way of framing these results is this: turnover on the nonwestern frontier side is lower, wherever there is a significant result. It is also lower for nonwesterners of all ages on the *native* side. But – it is *higher* for the two younger *native* age groups, on the native frontier side. This highlights the turnover difference for native versus nonwestern people.