Love thy neighbor?

An empirical test of neighborhood ethnicity change and Schelling behavior*

Jørgen Baun Høst†
University of Copenhagen
Department of Economics

April 6, 2025

Abstract

Does the ethnicity of your nearest neighbor affect your propensity to move? To provide causal evidence on this question, I use comprehensive administrative data with precise geospatial information and implement a nearest-neighbor research design that compares households within the same neighborhood who receive different-type neighbors as their nearest neighbors to ones who receive them "just down the road". I find asymmetry in residential responses based on neighbor ethnicity: Native Danish households increase their propensity to move within 2 years by approximately 1.6 percent compared to their baseline exit rate, when they receive a new non-Western neighbor. In contrast, non-Western households show no such response. This effect is primarily driven by low-SES native households responding to low-SES non-Western neighbors. These findings provide causal evidence of individually motivated segregation as theorized by Schelling (1971), though with more modest effects than documented in the United States.

Keywords: Schelling, segregation, ethnicity, neighborhood, KNN

JEL classification: J15, R23

*I am grateful for the guidance given by my supervisors, Andreas Bjerre-Nielsen and Nikolaj Arpe Harmon, in conjunction with valuable input by both Jeppe Søndergaard Johansen and Jan Moritz Johanning. I also thank the Data Science Lab team at Statistics Denmark for giving me access to the unique datasets used in this paper.

†In loving memory of Norredine Remaoun.

Contents

Character Count: $42,000$

1 Introduction

Residential segregation, the uneven distribution of ethnic groups within a confined geographic area, remains a persistent feature of Danish society, yet there is limited empirical evidence on how ethnic background directly influences residential sorting. To this end, I seek to answer the fundamental question: To what extent do households respond directly to the ethnic identity of their nearest neighbors when making the decision to move?

Schelling (1971) remains the foundational piece in the literature on theoretical determinants of segregation. The key prediction of Schelling (1971) is that neighborhoods may experience a rapid outflow of residents belonging to a majority, or "tip", when the share of (new) residents belonging to a minority reach a certain threshold. Schelling (1971) noted that this phenomenon may happen even if majority residents are relatively tolerant of other minorities.[1] This paper focuses on what Schelling called the "third" kind of segregation - individually motivated segregation - as distinct from organized segregation (such as historical Jim Crow laws) or economically induced segregation (clustering by income or education). I do not make the claim that organized or economically induced segregation do not exist or play a role in Denmark (perhaps less so than in the United States), but that individual preferences, even if it is not the main driver of observed segregation, still matters and deserves attention.

The Danish context offers a particularly valuable setting for this research question. Over the past four decades, Denmark has transformed from a relatively homogeneous society to one with increasing ethnic diversity, particularly in urban areas. This demographic shift, combined with Denmark's comprehensive administrative data infrastructure, provides an ideal environment to examine Schelling behavior in a European welfare state context.

Work by Schelling (1971) has spawned a rich literature on segregation from which many papers draw their inspiration from. Card et al. (2008) employ a single "tipping point" Schelling model in an RDD setting. They find discontinuity in the rate of change in minority shares ranging between 5 pct. (Portland, Oregon) to around 20 pct. (Los Angeles, California). Inspired by Card et al. (2008), Böhlmark and Willén (2020) estimate neighborhood tipping points in Sweden's largest cities (Stockholm, Malmo and Gothenburg) that range between 17-19 pct. minority shares. Blair (2017) further builds on Card et al. (2008) and the idea of neighborhood tipping points in the United States, but points to the

[1] To illustrate, I have simulated a simple version of Schelling's model, which takes inspiration from Mingarelli (2021). This can be found in Appendix C.

importance of outside options in this context. Blair (2017) argues that while some White American households may dislike neighbors from a different ethnic background, they may be constrained by unattractive alternatives. Here, outside option refers to the set of neighboring census tracts that exists within a household's Metropolitan Statistical Area (MSA).

While the tipping point literature has advanced our understanding of neighborhood-level dynamics, identifying the causal mechanisms underlying these patterns remains challenging. Bayer et al. (2022) address this challenge by developing a nearest neighbor research that examines household responses to specific neighbor changes at precise spatial location within a neighborhood. By comparing households within the same neighborhood who receive different-type neighbors either next-door or "just down the road", they isolate preferences about neighbors from preferences about neighborhood amenities. Once applied to a US setting, they find symmetric responses among both White and Black households, with 4-6 percent increases in baseline exit rates within 2 years following the arrival of a different-race neighbor. This paper adapts their methodological approach to a Danish context, allowing for a novel comparison of Schelling mechanisms across different institutional settings.

Identifying the causal role of preferences for neighborhood composition is vital for understanding persistent segregation. It has been widely established that neighborhoods, both who live there and amenities within, matter for long-run outcomes. For instance, Chetty et al. (2016) show the positive impact on earnings and college attendance when presented with the opportunity to move to a more affluent neighborhood. In a Danish context, Hasager and Jørgensen (2024) show the importance of neighborhood (both at the building and parish level) composition on long-term health outcomes using the quasi-experiment of the Danish Spatial Policy for refugee placement. Damm and Dustmann (2014) use the same policy as part of their empirical framework and finds that neighborhood crime rates is strongly associated with youth criminal behavior. In an American context, Caetano and Maheshri (2017) argue much of the residential sorting found in the US is the result of the parental preferences for the racial and socioeconomic composition of schools, which can lead schools to "tip" with school segregation even greater than neighborhood segregation. Bjerre-Nielsen and Gandil (2024) find similar patterns in a Danish context, where parental school preferences cause segregation in school attendance with a strong socioeconomic gradient in responses.

Perhaps most similar to me is Boje-Kovacs et al. (2024), who examine native flight in Denmark, both at the neighborhood and building level. They find that an increase in 30 percentage points in the share of non-Western foreigners in the neighborhood increase the propensity for natives to move out in any given year by around 8 pct.[2] Unlike these previous studies that often rely on neighborhood-level data, I leverage a unique administrative dataset that pinpoints individual residential locations with precise spatial coordinates and exact timing of moves. Using this granular data, I implement the nearest neighbor research design developed by Bayer et al. (2022), comparing households within the same neighborhood who receive different-type neighbors either next-door or just down the road. I focus specifically on interactions between native and non-Western households to test for potential asymmetry in Schelling behavior, extending previous work by examining these dynamics in the context of a European welfare state.

I find that native Danish household increase their propensity to move within 2 years by approximately 0.3 percentage points, an increase by 1.6 percent compared to their baseline exit rate, when they receive a new non-Western neighbor among their three nearest neighbors compared to "just down the road". This effect is robust across specifications and primarily driven by low-SES native households responding to low-SES non-Western neighbors. In contrast, non-Western households do not show the same type of behavior for new native neighbors, which highlights the asymmetry in Schelling behavior in Denmark.

The paper is structured as follows. Section 2 presents the theoretical framework and empirical strategy heavily influenced by Bayer et al. (2022). Section 3 describes the data sources, the steps taken to process the unique datasets used in this paper along with descriptive statistics on neighborhood composition. In section 4, I show my main results and heterogeneity in Schelling behavior. I conclude on my findings in section 5.

1.1 Definitions

In the paper, I define 3 mutually exclusive types of households. (i) *Native* households, where all members are of Danish origin; (ii) *non-Western* households, where at least one member is of non-Western origin and (iii) *Western* households, where at least one household member is non-Native and of Western origin, but with no members that have *non-Western* origin. I follow the definition of Western/non-Western countries by Statistics Denmark (2022).[3]

[2] Neighborhoods contain around 600 households in Boje-Kovacs et al. (2024). These are derived from work by Damm and Schultz-Nielsen (2008).

[3] Throughout the paper, I repeatedly write about "same"- and "different"-type neighbors. To be clear, same-type neighbors refers to two or more households that live in close proximity to each other and both fall within the same category defined above. If they differ in terms of household type, they are different-type neighbors.

2 Model

2.1 Empirical framework

I begin by formalizing the framework by Bayer et al. (2022) that models a household's decision to either move or stay in a neighborhood that evolves over time. I assume that the preferences of households are dynamic - they care about the quality and composition of their neighborhood currently, but also how they expect the neighborhood to evolve over time.

Consider the following utility function for an existing homeowner i that lives in neighborhood j with observable attributes Z_i that maps the dynamic binary choice of either staying or leaving in discrete time:

$$U_{i,j,t} = f(Z_{i,t}, X_{j,t}, \xi_{j,t}) + \sum_k g(Z_i, Z_{k,t}, D_{i,k}) + \epsilon_{i,j,t} \tag{1}$$

$f(\cdot)$ captures utility derived from the both the observed $X_{j,t}$ and unobserved $\xi_{j,t}$ amenities of a neighborhood. $g(\cdot)$ captures the utility derived from the characteristics of each neighbor k, who lives at a distance of $D_{i,k}$ away. The distance parameter $D_{i,k}$ in the utility function captures the spatial dimension of neighbor effects, allowing the model to account for the intuitive concept that interactions with immediate neighbors may be more consequential than with those living further away. This spatial dimension is central to the nearest neighbor research design. $\epsilon_{i,j,t}$ captures the idiosyncratic taste of household i's specific home.

Clearly, unobserved neighborhood amenities constitute a problem to identification. Should you be so lucky to have access to data on neighborhood amenities at a granular scale, you could include these in a regression model, but they may be endogenously determined themselves. For instance, school quality is correlated with neighborhood income, but what made them good in the first place? To further illustrate the problems involved unobserved amenities, suppose a household faces school closure in their local area, but at the same time get a new neighbor immediately next door that they find unpleasant. Is the decision to move then the result of the school closure or the new neighbor? There is no clear answer to this and illustrates why these effects are difficult to disentangle.

Furthermore, the dynamic nature of households preferences itself also pose a threat to the identification of Schelling behavior. Consider the following Bellman equation that recursively maps utility over time:

$$V_{i,j,t} = f(Z_{i,t}, X_{j,t}, \xi_{j,t}, \alpha) + \sum_k g(Z_{i,t}, Z_{k,t}, D_{i,k}, \beta) + \delta \mathbf{E}[V_{i,j,t+1}] + \epsilon_{i,j,t} \qquad (2)$$

Where δ is a discount rate. The second part of the Bellman equation highlights the issue of expectation for the future development of the neighborhood. If a household gets a new neighbor with a different ethnicity, do incumbent households respond directly to the ethnicity of their neighbor or is the new neighbor a signal of future neighborhood development that may or may not be to their taste?

This framework formalizes the intuition behind Schelling's classic model of segregation Schelling (1971) but extends it by explicitly accounting for both the spatial dimension of neighbor interactions and the forward-looking nature of household decision-making. The challenge is to empirically separate direct preferences over neighbor attributes from responses to correlated neighborhood characteristics or expectations.

2.2 Empirical strategy

To account for the issues above and ensure credible identification, I compare households who receive a new different-type neighbors among their *nearest* neighbors to those who receive a new different-type neighbor "just down the road". I denote these "treated" and "control" households. The difference in moving propensity between the two is an estimate of Schelling behavior. Intuitively, this makes sense, because i) treated and control households live in the same neighborhood and thus experience (practically) the same (un)pleasantries of the local area and ii) they experience (almost) the same signal of future neighborhood development. I formalize this intuition in more detail below while following Bayer et al. (2022).

Consider a home i. The closest neighbors to i are homes $j \neq i$ have ordinal rank K in distance from i. In this paper, I focus on $K = 40$ nearest neighbors.[4]

I define nearest neighbors as those living at rank $k_{nearest} \in \{1, 2, 3\}$ and those "just down the road" as $k_{near} \in \{4, 5, 6\}$ and close neighbors as $k_{close} \in \{7, ..., 40\}$.

With equation 1 in mind and abstracting from the time t indices, consider the moving propensity Y_i of household i in response to a new neighbor of different ethnicity type e' that moves in a home that is at ordinal rank K distance away:

[4]The choice is in part due to computational constraints, as datasets on this form quickly explode in size. Further, I do this to compare results with Bayer et al. (2022) who narrows the scope at the same scale.

$$Y_i(e', k) = \mathcal{P}[e', k] + \xi_i B(e', k) + \rho_i + \omega_j \tag{3}$$

The first term denotes the ’tipping point’ in the spirit of Schelling (1971), ie. the direct preference $\mathcal{P}$ for living close to a k-nearest neighbor of different ethnicity.[5] The second term denotes the difference in future amenities related to the arrival of a new different-type K-nearest neighbor. Finally, ρ_i and ω_j captures the idiosyncratic factors which affect household i’s moving propensity and neighborhood j as a whole.

I am specifically interested in $\mathcal{P}[e', k_{near}]$, ie. the moving propensity in response to a new different-type neighbor among your *nearest* neighbors. To do so, I difference the moving propensity between those who get a new different-type neighbor a among your $k_{nearest}$ to those who get a new different-type neighbor b slightly further away, k_{near} in a neighborhood j:

$$\begin{aligned} Y_a(e', k_{nearest}) - Y_b(e', k_{near}) = {} & (\mathcal{P}[e', k_{nearest}] - \mathcal{P}[e', k_{near}]) \\ & + (\xi_a B(e', k_{nearest}) - \xi_b B(e', k_{near})) \\ & + (\rho_a - \rho_b) + (\omega_j - \omega_j) \end{aligned} \tag{4}$$

Assumption 1 $\mathcal{P}[e', k_{nearest}] - \mathcal{P}[e', k_{near}] > 0$

Formally, this implies that the effect of getting a new different-type neighbor among your, say, three nearest neighbors is greater than it is among your neighbors of rank four to six. This implies that any effect I may find is likely to be a lower bound of the effect in getting a new different-type neighbor. Conversely, like Bayer et al. (2022) I find the assumption $\mathcal{P}[e', k_{near}] = 0$ too strong to make. This assumption implies zero effect on the moving propensity of getting a new different-type neighbor just outside your nearest neighbors.

Assumption 2 $\xi_a B(e', k_{nearest}) - \xi_b B(e', k_{near}) \approx 0$

This assumption implies that the difference in expectations of future amenities following a new different-type neighbor among your nearest and near neighbors is (near) zero. This assumption is likely to hold as households that live in close proximity essentially live in the same environment - they have similar access to neighborhood amenities, like public transportation and schools. If this assumption is violated, then estimates of Schelling behavior would also capture the differential in expectations. This problem could

[5]Note, that this is not a direct estimation "tipping point" of Schelling (1971). To my knowledge, the closest to this is work by Böhlmark and Willén (2020) who estimate city-level tipping points in Sweden.

arise, if households with a new different-type neighbor among their nearest neighbor systematically formed more pessimistic expectations about the development of neighborhood amenities than those with a new different-type neighbor slightly further away.

Assumptions 1 and 2 leaves:

$$Y_a(e', k_{nearest}) - Y_b(e', k_{near}) = \mathcal{P}[e', k_{nearest}]^* + \rho_a - \rho_b \tag{5}$$

Assumption 3 $\mathcal{P}[e', k_{nearest}]^* \perp \rho_a - \rho_b$

Assumption 3 states that new neighbors are quasi-randomly assigned within the local neighborhood j from the perspective of incumbent households. In other words, the specific home where a new neighbor moves in is not systematically related to the existing homeowner's individual preferences. While households may select into specific neighborhoods based on broader preferences, the precise location where a housing unit becomes available - and subsequently occupied by a household of ethnicity e - is unlikely to be systematically related to the unobserved characteristics of existing households living a few doors away. It should be noted here that because I am comparing two or more households within a specific neighborhood, I implicitly assume that they have the same "baseline" moving propensity prior to a new different-type neighbor moving in.

Averaging over J neighborhoods yields a consistent estimate of the average treatment on the treated (ATT) $\overline{Y(e', k_{nearest}) - Y(e', k_{near})}$ conditional on observable characteristics. I empirically formalize this approach below:

$$\begin{aligned} Y_{i,j,t} =& \beta_1 \mathbb{I}[e', k = n_{nearest}] + \beta_2 \mathbb{I}[e', k = n_{near}] + \beta_3 \mathbb{I}[e', k = n_{close}] \\ &+ \gamma Z_{i,j,t} + \omega_{j,t} + \epsilon_{i,j,t} \end{aligned} \tag{6}$$

$Y_{i,j,t}$ denotes the outcome of interest, an indicator ($\times 100$) for whether household i moves within 2 years following a new different-type neighbor.

I include household-level $Z_{i,j,t}$ control variables that affects the propensity to move. These include household age, household size, household income and tenure at address (how long you have lived at your current address at a moment in time).

The parameter(s) of interest is $\beta_1 - \beta_2$, which represents the difference in moving propensity in response to a new different-type neighbor. I choose to segment and add treatments in "bins" of ordinal ranks K. Specifically, I define $k_{nearest} \in \{1, 2, 3\}$, $k_{near} \in \{4, 5, 6\}$ and $k_{close} \in \{[7-10], [11-20], [21-30], [31-40]\}$. First and foremost, the reason I do this is to increase precision by adding "treatments", but it also allows me to examine whether the effect of receiving a different-type neighbor decays with distance. It also facilitates direct comparison with Bayer et al. (2022).

The inclusion of neighborhood-by-quarter fixed effects $\omega_{j,t}$ is central to my identification strategy. These fixed effects ensure I compare only households within the same neighborhood in the same time period, thereby controlling for all time-varying neighborhood characteristics that might simultaneously affect both the arrival of different-type neighbors and moving decisions.

I choose to focus on quarter-level time frequency. First, my prior of the outcome of moving is that it may be rather sparse in some places. Thus, I need to allow household to update their beliefs about the attractiveness of their current neighbors and neighborhood. Second, I believe that yearly time frequency do not capture salient neighborhood development.

3 Data

This section describes the data sources and methodology used to construct the household-level dataset that forms the foundation of my analysis. I focus on the time period from 1985-2020, which is the longest period my administrative data sources covers. [6]

3.1 Administrative data sources

My analysis draws on comprehensive administrative microdata from Statistics Denmark, which can be linked at the individual level. The dataset integrates information from multiple administrative registers, providing detailed demographic, socioeconomic, and geographic information for the entire Danish population. Administrative data tables are denoted **TABLE_NAME** and corresponding variables of interest, **variable_name**.

Specifically, I combine the following core registers:

- Population register (**BEF**): Contains fundamental demographic variables including birth date (**foed_dag**), family structure through parental identifiers (**mor_id**, **far_id**), country of origin (**opr_land**), and partner identifiers (**aegte_pid**/**e_faelle_pid**).
- Income register (**IND**): Provides detailed income data, from which I extract gross income (**perindkialt_13**)—encompassing both wage earnings and public transfers—and net wealth excluding pension assets (**form**/**formrest_ny05**).
- Labor market register (**RAS**): Records labor market attachment, from which I obtain employment status (**beskst13**).
- Education register (**UDDF**): Documents educational attainment, from which I extract highest completed education level (**hfaudd**), categorized according to the DDU nomenclature and converted to education length in years.[7]

3.2 Demographic context

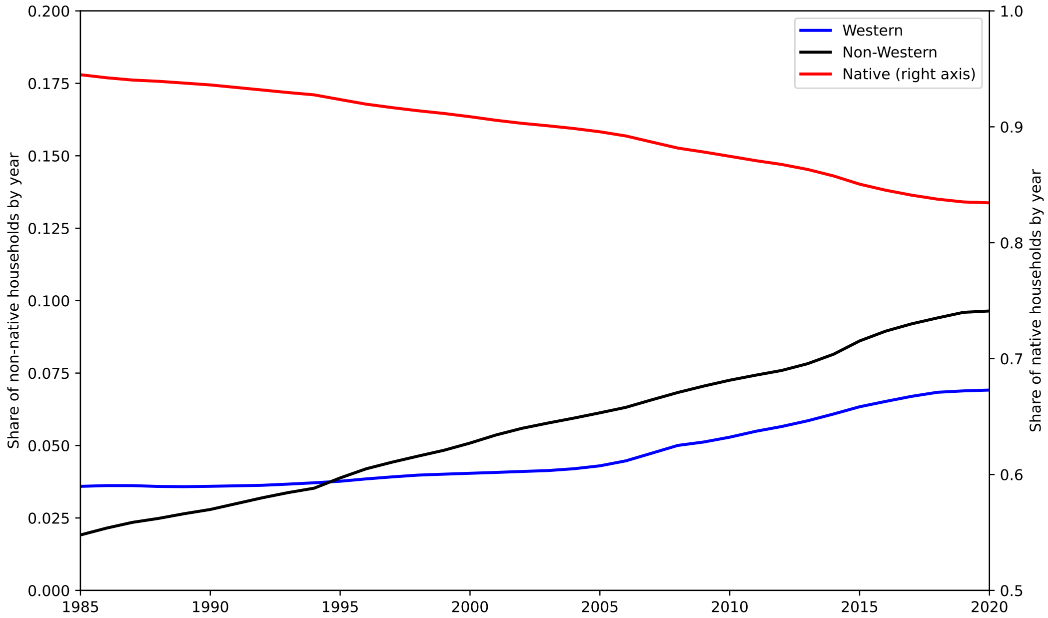
The period of my analysis coincides with a significant demographic transformation in Denmark, which provides an ideal setting to study Schelling behavior. Figure 1 illustrates this transformation: the share of non-Western households grew from approximately 2 percent in 1985 to around 10 percent by 2020. This substantial change creates natural variation

[6] All code is available here. The data processing relies extensively on Polars (Vink et al. (2025); Eechoud (2024)), a high-performance DataFrame library that enables efficient parallel processing of datasets containing several billion rows. I am grateful to the developers for creating this exceptional tool.

[7] Exact categorization of completed education length in years can be found in table A.8. For details on the DDU nomenclature, see Statistics Denmark (2024a).

in the frequency with which native and non-Western households encounter different-type neighbors.

Figure 1: Distribution of households (1985-2020)



Note: This is based on own calculations and shows the distribution of households by type as defined in section 1.1. I sample households that are present at December 31st in each year from 1985.

3.3 Geospatial

The main dataset I use is **BOPAEL_KOORD**. This dataset contains all historic addresses in four dimensions:

$$\mathbf{p} = (x_E, y_N, z_F, z_D) \tag{7}$$

Where $\mathbf{p}$ is a tuple of x/y/z-coordinates, where subscripts denote east, north, floor and door, respectively, in the ETRS89-projection. This projection is specifically tailored to Northern Europe such that distances between coordinates, whether these are Manhattan or euclidean, are measured in meters. Crucially, the dataset contain both the start and end date of residency at those coordinates. These are derived from administrative data, where individuals report their main address. Thus, in effect, the outcome of "moving" is synonymous with a "change in address".

While the x_E, y_N and z_F-dimension are all numerically represented, the z_D-dimension is not.[8] For instance, a sidedoor can be represented by a string such as *"TV"*, if the entrance

[8]To be clear, the floors are represented as the following. Ground floor is assigned the value of 0, first floor the value of 1 and so forth.

door to your apartment is to the left. Similarly, sidedoor can also contain lettering *"A", "B"* and so forth. To facilitate precise identification of nearest neighbors, I choose to assign the strings *"TV"*, *"MF"* and *"TH"* a number of -1, 0 and 1. Additionally, any additional lettering are assigned in alphabetic order, such that *"A"* is assigned the value 1, *"B"* the value of 2 and so forth. This ensures that I correctly rank nearest neighbors in dense neighborhoods.[9]

A key component of my identification strategy is that I specify which *neighborhood* a household lives in at a given point in time. This is parameterized by $\omega_{j,t}$ in equation 6. There exists no clear consensus of what constitutes a neighborhood in the geographical sense. For instance, I could use administrative borders such as municipalities or parishes. However, I believe there is a number of issues by doing this. First, municipalities are way too large to elicit salient social interactions and, second, I believe parishes to have borders that are too inconsistently drawn.[10] I rely on work by Bøje-Kovács et al. (2023), who use the Danish Kvadratnet, a graphical representation of Denmark as squares down to the 100m-by-100m level, to delineate "fixed" neighborhoods as polygons that contain at least 100 inhabitants. To do this, they apply a heuristic version of the MaxP-regionalization algorithm by Wei et al. (2021). Because I fear that the outcome of moving in may be relative sparse in some areas, I elect to further aggregate this to the level of a minimum 500 inhabitants to ensure sufficient statistical power while maintaining meaningful social interaction boundaries. This leaves me with a set of 5,933 unique neighborhoods across Denmark.

3.4 Nearest neighbors

With the definition of households and geospatial dataset in mind, I now describe the approach for identifying nearest neighbors. From $\mathbf{p}$, I construct a KD-tree (Bentley (1975)) containing all unique addresses. The KD-tree efficiently partitions a k-dimensional space to enable fast spatial queries. In my application, $k = 4$, corresponding to the 4-dimensional point representation in $\mathbf{p}$.

The KD-tree construction follows a recursive process where the dataset is split along alternating dimensions at each level. At each node, the data is partitioned using the median value along the current dimension, creating a balanced tree structure. For any query point $q \in \mathbb{R}^4$, I can efficiently retrieve either the K nearest neighbors (by ordinal rank), which is what I focus on, or all neighbors within a specified radius r. The nearest neighbor

[9] If addresses contain no floor or sidedoor information (a single-family villa for example), these are assigned the value 0 and fed to the KD-tree. Otherwise, distance calculation for these addresses would be impossible given the "missing" data.

[10] In 2024, the sizes of parishes ranged from less than 100 to over 20,000 inhabitants (Statistics Denmark (2024c)).

query procedure involves traversing the tree to find the K nearest points according to the Euclidean (L_2-norm) distance metric:

$$d(p_i, p_j) = \sqrt{\sum_{d=1}^{4}(p_i^d - p_j^d)^2} \tag{8}$$

where p_i^d represents the d-th dimension of point p_i. Because not all addresses exist for the same amount of time (new homes are built or old ones torn down), I make sure to "buffer" extra addresses by widening my search space.

3.5 Households

Given the richness of the data, I need to carefully define what constitutes a "household". To do this, I borrow from the graph theory literature. Define individuals as nodes and spatio-temporal overlap as edges. Edges exist only if people overlap in time and space. For N number of sequences, define the $N \times N$ adjacency matrix $\mathbf{A}_{i,j}$:

$$\mathbf{A}_{i,j} = \begin{cases} w_{i,j} & \text{if people overlap in time and space} \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

Weights are determined by:

$$w_{i,j} = \frac{\max(0, \min(t_{i,end}, t_{j,end}) - \max(t_{i,start} - t_{j,start}))}{t_{i,end} - t_{i,start}} + \mathbb{I}[i = j = G] \tag{10}$$

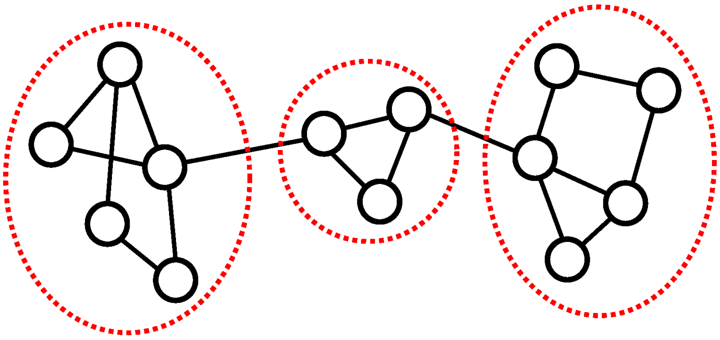
The first part of equation 10 expresses how much sequence i overlapped with sequence j as a fraction of duration of sequence i. The second part is an indicator function to flag if the two sequences are part of the same group G, be it as family or as partners.

Consider the fact that not all members of a household may move out around the same time, which further complicates credible identification of "Schelling behavior". For example, suppose person A and person B moves in at the same time in a home. Person A moves out just before B, but a person C moves into the home and lives with person B for a short amount of time, before B moves out. Linking the sequences together, or in graph theory terms identifying *connected components*, would yield a single household of A, B and C, when it most intuitively makes sense to have A and B as one household and C as another.

To formalize this approach, I use administrative registers as an intermediate to evaluate different "household partitions", specifically the variable **familie_id**. This group variable identifies individual who live under the same roof, but the variable changes value changes if, say, marital status changes or a child moves from their parents.[11]

Maximizing edge weights defined in equation 10, scaled by the size of the household partition, is what I define a "household" as. Consider Figure 2 as an illustration of this approach. Maximizing edge weights would yield three different "stable" households instead of one. I drop sequences that started before 1960, when I explore the **BOPAEL_KOORD** dataset. I find that sequences prior to this are too sparse or start at arbitrary dates like 1st January 1900. In total, the process yields around 14,000,000 historic household *sequences*.

Figure 2: Household definition



[11]For more detail, see Statistics Denmark (2024d).

3.6 Estimation sample

I impose a number of restrictions on the sample of households to build my estimation sample:

1. **Income/wealth**: I require households to have gross real yearly equivalised incomes between 200,000 DKK and 1,000,000 DKK. In similar vein, I keep household that have real equivalised net wealth (excluding pensions) between -200,000 DKK and 750,000 DKK. This is to ensure similar residential mobility across households.

2. **Age**: I further require that the oldest member of the household is between 30 and 60. These restrictions weeds out instances where members of households may be studying or retired, as I believe both of these groups to have significant differences in residential mobility.

3. **Ethnicity**: I restrict my analysis to native and non-Western households, following the definition set out in section 1.1. This choice is mainly driven by the demographic development that coincided with the time period I study. Furthermore, previous evidence (Böhlmark and Willén (2020), Grand and Szulkin (2002), and Boje-Kovacs et al. (2024)) mainly points to anti-immigration sentiments between native households and non-Western immigrants in a Scandinavian context.

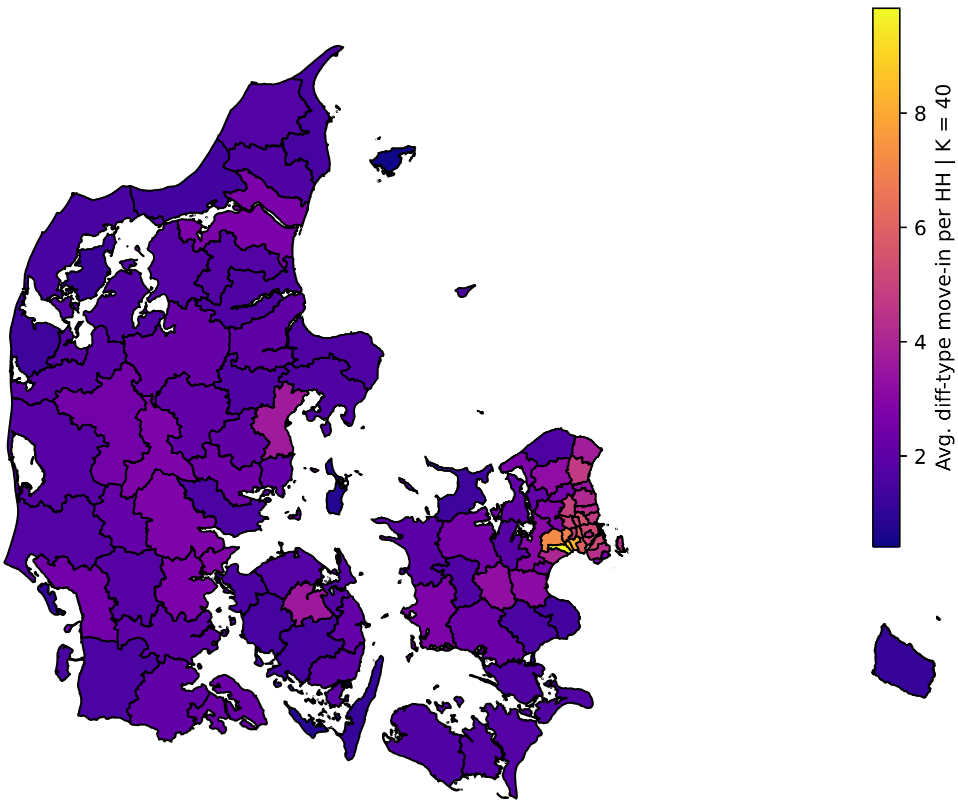
4. **Proximity**: I require that the nearest neighbor to a household is within 25 meters. This is to ensure that social interactions between new different-type neighbors are more likely to take place. I present evidence of why this assumption is reasonable in Table A.9, as household have limited response to a new different-type nearest neighbor beyond 25 meters. Results from table A.9 suggests that I could further limit this distance, but with the potential inaccuracies in the procedure of assigning strings with a numerical value, I elect to be "conservative" and set the boundary to 25 meters. Furthermore, I require that neighborhoods to have between 1,000 and 25,000 people per square kilometer to ensure homogeneity across different types of neighborhoods. This leaves me with 3,451 unique neighborhoods.

3.7 Summary statistics and spatial patterns

Figure 3 depicts the incidence of new different-type neighbors among your $K = 40$ nearest neighbors for native households split by municipality since 1985. The color scheme is chosen to highlight the intensity of new different-type neighbors that native households receives during their residence. Darker blue/purple colors represent "low" intensity with orange/yellow representing "high" intensity.

Perhaps the most striking pattern is the east-west and urban-rural divide in the incidence of new non-Western households clearly visible in the figure. Apart from Aarhus and Odense, the vast majority of new non-Western neighbors are concentrated in Copenhagen and its surrounding municipalities. Perhaps a little surprising at first, it is, however, not the municipality of Copenhagen with the highest incidence (I show below why this is the case). It is in fact in Ishøj, where the average native household experience over 9 new different-type neighbors among their $K = 40$ nearest neighbors. In comparison, native households in Copenhagen get around 6 new non-Western neighbors during their residence with around 4 for native households in Aarhus and Odense.

Figure 3: Incidence of new different-type neighbors (1985-2020)

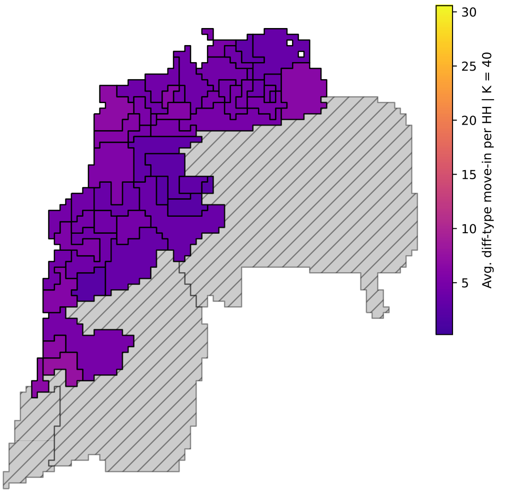


Note: The figure show the variation in receiving a new-non Western neighbors within the 40 closest parcels for native households. Municipal borders correspond to the ones imposed by "Kommunalreformen" in 2007. Household types are split up in to three types, see section 1.1 for more details.

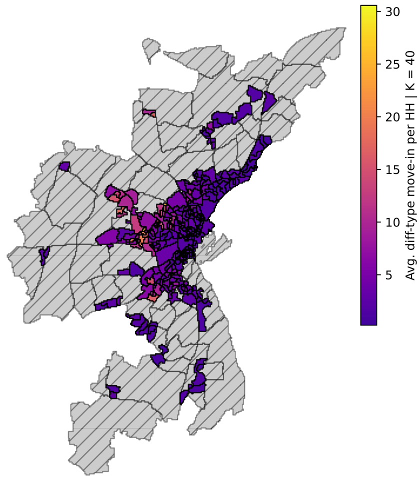
This pattern makes sense in a historical context. While the Danish government did implement a spatial refugee dispersal policy which successfully allocated refugees equally across municipalities (Hasager and Jørgensen (2024)), the majority of immigrants has historically come under family reunification terms (Statistics Denmark (2024b)) which led to clustering to in major urban areas.

Figure 4 shows the incidence of new different-type neighbors across neighborhoods in three different cities/municipalities since 1985 as defined in section 3.3. Residents in some neighborhoods in Copenhagen get upwards of 30 new non-West neighbors during their residence. Specifically, these are neighborhoods that are characterized by high-concentration of public housing, such as Mjølnerparken in Nørrebro. Other neighborhoods within Copenhagen get less than 2 new non-Western neighbors - these are considered to be moer affluent neighborhoods. The same pattern is prevalent for Aarhus, where the average native resident in some neighborhoods (like around Gellerupparken) get more than 15 new non-Western neighbors.

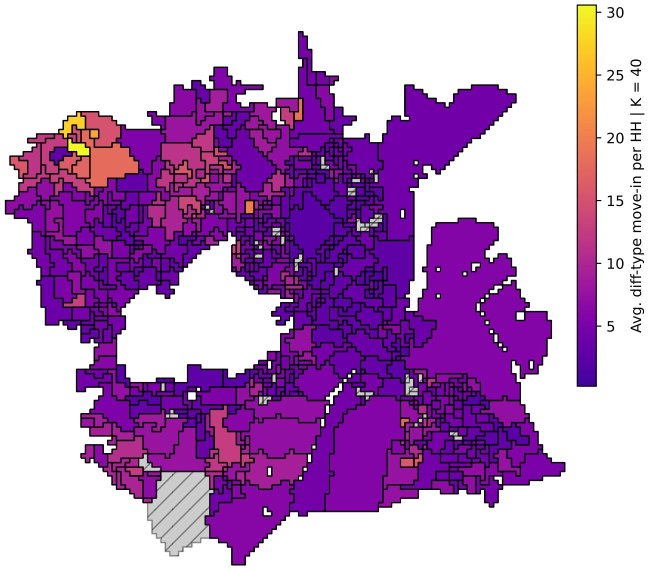
Figure 4: Incidence of new different-type neighbors at the neighborhood level (1985-2020)



(a) Ishøj



(b) Aarhus



(c) Copenhagen

Note: The figure show the variation in receiving a new-non West neighbor for native households at the *neighborhood* scale for three different municipalities. Neighborhoods not in the sample are greyed out. Household types are split up in three types (native/non-West/West), see section 1.1 for more details. Neighborhoods are defined in section 3.3. See Figure B.1 for an unconstrained version of this figure.

Table 1 presents summary statistics for the two samples of Native and non-Western households as defined in section 3.6. Columns *All* describe each quarter-by-year observations for a given household where all covariates are observed in the period between 1985 and 2020 and where household may be "at-risk" of receiving a new different-type neighbor. Columns *Nearest* and *Close* each describe the year-by-quarter observations, where a household experience a new different-type neighbor either as their nearest neighbor ($K \in [1, 2, 3]$) or as their "close" neighbor ($K \in [4, 5, ..., 40]$).

Table 1 reveals an intriguing pattern in the first row across both panels. "Treated" households (both native and non-Western) demonstrate a notably higher likelihood of relocating within two years after gaining a different-type neighbor: 23-24 percent compared to 19-20 percent for their "control" counterparts. However, a critical consideration for this analysis is that this apparent Schelling response could be influenced by various confounding factors, including changes in neighborhood quality or selection effects. This observation underscores the value of employing the nearest neighbor research design, which helps address these potential confounders.

The following rows reveal that treated native households are considerably less wealthy than the full sample of native households (48,500 DKK vs 81,000 DKK) in addition to earning less (337,000 DKK vs 344,000 DKK), but that this difference, though still relatively large, is less between the "treated" and "control" native households. This pattern suggests that receiving a new non-Western neighbor is not random and is more likely to happen in neighborhoods that are already experiencing demographic change than in more affluent areas. Furthermore, this difference may also reflect micro-sorting by wealth even within neighborhoods, which further underscores the importance of the within-neighborhood research design. Treated non-Western households have slightly lower net wealth (41,000 DKK) than control non-Western households (47,000 DKK) with the full sample falling between these values. Unlike for native households, there seems to be no sorting based on the wealth gradient alone.

Interestingly, non-Western households in the sample are on average better educated (defined as the person in the households with the longest education) than native households. On average, the difference between them are about 2 years of education for the longest educated household member. This is likely to be reflected in their preferences for neighborhood composition and, thus, response to receiving new native neighbors.

Table 1 also highlights some striking patterns regarding the neighborhood characteristics. First, the neighborhoods in which these experiments happen tend to be concentrated in relatively dense and more mixed neighborhoods in line with figure 4. For instance, non-Western households as a whole live in considerably denser neighborhoods than their native counterparts. Second, and perhaps an indication of Schelling behavior, native households tend to live in less dense, more affluent and less-integrated neighborhoods. For instance, treated native households live in much denser neighborhoods with almost double the share of non-Western households (15 percent) compared to the whole sample of native households (8 percent). It is therefore not surprising to see that this is also reflected in the average distance to new different-type neighbors. For both samples, these tend to, on average, have new different-type neighbors that live 30-33 meters for their $K = 40$-nearest neighbors.[12]

[12]Keep in mind the potential inaccuracies when calculating the euclidean distance, see section 3.3 for more details.

Table 1: Summary statistics

	Native households						Non-Western households					
	All		Nearest		Close		All		Nearest		Close	
Household characteristics												
Move within 2 years	17.45	(37.95)	23.19	(42.21)	19.74	(39.80)	19.14	(39.34)	24.01	(42.71)	19.34	(39.50)
Real inc. (1000s) DKK	344.11	(120.22)	337.16	(119.58)	341.07	(120.29)	313.13	(111.14)	317.37	(114.96)	314.62	(112.27)
Real net wealth (1000s) DKK	81.03	(202.68)	48.51	(185.87)	62.08	(193.44)	45.52	(163.75)	41.34	(160.52)	46.95	(165.29)
Employed	0.86	(0.34)	0.83	(0.37)	0.84	(0.37)	0.83	(0.38)	0.84	(0.37)	0.83	(0.37)
Years of education	10.15	(6.24)	10.03	(6.26)	10.13	(6.25)	12.12	(5.34)	12.13	(5.47)	12.10	(5.37)
Distance to neighbor			3.20	(4.29)	33.15	(20.69)			2.81	(3.70)	30.60	(18.15)
Household size	1.81	(1.06)	1.63	(0.94)	1.70	(0.99)	2.14	(1.29)	1.96	(1.20)	2.13	(1.27)
Oldest household member	44.25	(8.93)	43.51	(9.03)	44.18	(9.00)	43.37	(8.71)	41.84	(8.61)	43.26	(8.69)
Neighborhood characteristics												
Population density	7418.23	(7644.95)	8969.51	(8151.09)	8387.60	(7945.02)	9198.14	(8155.82)	9667.16	(8521.05)	9331.84	(8326.29)
Native share	0.89	(0.10)	0.81	(0.14)	0.83	(0.13)	0.76	(0.18)	0.78	(0.16)	0.78	(0.16)
Non-Western share	0.08	(0.09)	0.15	(0.14)	0.13	(0.12)	0.19	(0.17)	0.17	(0.15)	0.18	(0.16)
Real income (median)	322.85	(32.94)	318.41	(35.49)	320.60	(34.61)	317.40	(35.96)	319.73	(36.69)	318.51	(35.72)
Real net wealth (median)	42.74	(55.89)	27.58	(44.34)	32.01	(48.11)	27.87	(44.55)	27.54	(42.75)	27.90	(43.95)
N	33,496,551		761,842		4,671,062		3,330,242		414,103		1,443,373	

Note: This table shows presents summary statistics for households "at-risk" of receiving a different-type neighbor. Standard deviations in parenthesis. Income and wealth are equivalised to facilitate comparison between households of different size and composition. The *All* column denotes quarter-by-year observation for the sample of household defined in section 3.6. The *Nearest* ("treated") and *Close* ("control") columns denote instances, where a household experienced a new different-type among their $K \in [1, 2, 3]$ nearest neighbors or close neighbors ($K \in [4, 5, ..., 40]$).

3.8 Residential sorting over time

To examine trends in residential sorting, I show the development of the proportion of same-type neighbors from 1990-2020 in figure 5, ranging from the $K = 5$ to $K = 100$ nearest neighbors. For native households, the proportion with exclusively native nearest neighbors has remained high and even increased over time, particularly at larger distance bands. By 2020, approximately 60 percent of native households had between 80-100 same-type neighbors among their 100 nearest neighbor. In 1990, this share was only around 40 percent.

To this end, I have included a quick and dirty "counterfactual" simulation that randomly assigns the type of households with conditional based on the 1990 distribution of household types. Let $C_i \in \{\text{native}, \text{non-Western}, \text{Western}\}$ represent the type of the i-th household. For the counterfactual scenario, I maintain the 1990 distribution of household types:

$$\Pr(C_i = c) = \mathrm{p}_{c,1990}, \quad \text{for } c \in \{\text{native}, \text{non-Western}, \text{Western}\}$$

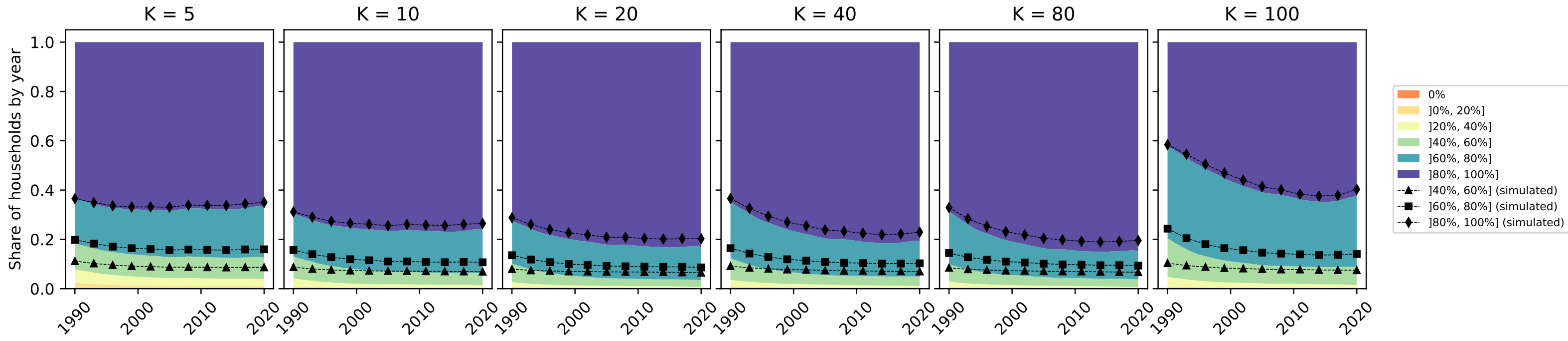
where $p_{c,1990}$ is the proportion of households of type c in the year 1990. We draw household types $\{C_1, C_2, \ldots, C_N\}$ from this categorical distribution while maintaining their spatial locations fixed. The resulting share of same-type neighbors S_i^K for each household is then computed based on these counterfactual type assignments:

$$\begin{aligned}
\mathrm{p}_{\text{native},1990} &= 0.95 \\
\mathrm{p}_{\text{non-Western},1990} &= 0.02 \\
\mathrm{p}_{\text{Western},1990} &= 0.03
\end{aligned}$$

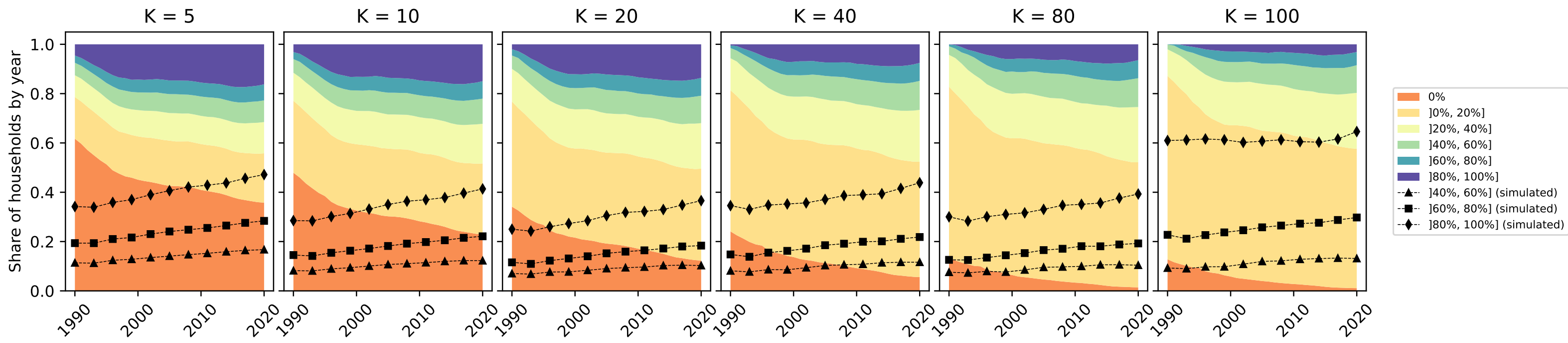
The counterfactual (shown as dashed lines in figure 5) suggests that the increasing segregation observed for native households exceeds what would have been expected, had the distribution of household types been unchanged since 1990. This constitutes preliminary indications of Schelling behavior.

The picture is different for non-Western households. I interpret the upward trends in all of the distance bands neighborhoods as indications of a trend opposite trend to that of native households. Holding the 1990 distribution of household types fixed, these quintiles would narrow, thus indicating a pattern of integration and not segregation.

Figure 5: Same type neighbor by K-nearest proximity (1990-2020)



(a) Native households



(b) Non-Western households

Note: This figure shows how the neighbor composition has changed over time for native and non-Western households, respectively. To construct this figure, I sampled which ever households and their neighbors were present at December 31st for each year between 1990 and 2020. I choose 1990 to not clutter the axis too much. Going from left-to-right, it shows the share of same type neighbors starting from your $K = 5$ up to $K = 100$ nearest neighbors. The dashed lines are "counterfactual" simulations of the top three quintile percentage bins holding the 1990 distribution of household types fixed.

3.9 Balance test

Before I present my main results of equation 6 and the parameters of interest, $\beta_1 - \beta_2$, I formally conduct a series of "balance" tests to check for any (mean) differences in observable characteristics between treatment and control households. These balance bests compares households that receive a new different-type neighbor among their nearest neighbors ($K = [1-3]$) compared to those "just down the road" ($K = [4-6]$) within the same quarter. If arrival of new neighbors is indeed quasi-random at this granular geographic scale, we should observe no systematic difference in observable household characteristics between these groups. To do this, consider the following equation, almost identical to equation 6:

$$X_{i,j,t} = \phi_1 \mathbb{I}[e', k = n_{nearest}] + \phi_2 \mathbb{I}[e', k = n_{near}] + \phi_3 \mathbb{I}[e', k = n_{close}] + \omega_{j,t} + \epsilon_{i,j,t} \quad (11)$$

Where $\mathbf{X_{i,j,t}}$ are observables at the household level. The coefficient of interest is $\phi_1 - \phi_2$. These include real equivalised income, real equivalised net wealth, oldest household member, tenure, employment status, educational length and household size. Employment is defined as at least one working household member in a given year. Education length is defined as the best educated household member measured in years. I choose to cluster standard errors at the neighborhood level, which corresponds to the level at which the key variation in my analysis occurs, see figure 4.

Starting with Table **??**, the first column show statistically significant, though economically relatively small differences in income between treated and control native households. The relative difference in net wealth between treated and control households is relatively larger, which I acknowledge is a threat to identification given that this offers greater mobility. However, the remaining columns indicate relatively small differences. The oldest member of treated households are only 0.04 years younger. Furthermore, there is only a mean difference in around 2 days in tenure between control and treatment households in conjunction with almost the same employment rate, education length and household size.

Table 3 reports corresponding estimates for non-Western households. Interestingly, the mean difference along the income and wealth dimension between treated and control non-Western households are economically smaller in scale and statistically indistinguishable from zero. The same goes for the difference in employment rates, household size and education length between the two. Treated non-Western households are on average 0.06

years younger, are similar in size and have difference in tenure length of around 1 month.

Table 2: Balance test (native)

	Income (1,000)	Net wealth (1,000)	Oldest HH member (years)	Tenure (days)	Employed	Educ. length (years)	HH size
New diff neighbor $k_{nearest}$ v k_{near}	-1.408***	-1.903***	-0.043*	1.936	-0.002**	-0.035*	-0.009***
	(0.254)	(0.402)	(0.020)	(5.999)	(0.001)	(0.014)	(0.002)
N	5,365,811	5,365,811	5,365,811	5,365,811	5,365,811	5,365,811	5,365,811
Neighborhood-by-quarter FE	X	X	X	X	X	X	X
Mean of dependent variable	341.26	60.61	44.05	2646.87	0.84	10.11	1.70
Number of neighborhoods	3444	3444	3444	3444	3444	3444	3444

* p < 0.05, ** p < 0.01, *** p < 0.001 Standard errors (in parenthesis) are clustered at the neighborhood level. The table reports the estimate of $\phi_1 - \phi_2$ from equation 11. Table A.1 contains the complete set of coefficients.

Table 3: Balance test (non-Western)

	Income (1,000)	Net wealth (1,000)	Oldest HH member (years)	Tenure (days)	Employed	Educ. length (years)	HH size
New diff neighbor $k_{nearest}$ v k_{near}	-0.065	0.171	-0.061*	31.109***	0.001	-0.011	0.001
	(0.347)	(0.522)	(0.028)	(7.656)	(0.001)	(0.018)	(0.004)
N	1,795,109	1,795,109	1,795,109	1,795,109	1,795,109	1,795,109	1,795,109
Neighborhood-by-quarter FE	X	X	X	X	X	X	X
Mean of dependent variable	315.64	46.50	42.94	2298.84	0.84	12.11	2.11
Number of neighborhoods	3332	3332	3332	3332	3332	3332	3332

* p < 0.05, ** p < 0.01, *** p < 0.001 Standard errors (in parenthesis) are clustered at the neighborhood level. The table reports the estimate of $\phi_1 - \phi_2$ from equation 11. Table A.1 contains the complete set of coefficients.

While the balance tests reveal some statistically differences between treatment and control native households in income and wealth, the nearest neighbor research design by Bayer et al. (2022) offers a great approximation of a quasi-experimental design. In the next section, I present my main findings of Schelling behavior.

4 Main results

This section presents estimates of Schelling behavior among native and non-Western households in Denmark. I find evidence that the propensity to move for native households increases with the arrival of new different-type household among their nearest neighbors rather than slightly further away, consistent with the prediction of the segregation model by Schelling (1971).

Table 4 and 5 report estimates $\beta_1 - \beta_2$ from equation 6 as proxy for Schelling behavior. All columns include neighborhood-by-quarter fixed effects ($\omega_{j,t}$ from equation 6) to facilitate comparable analysis. The first column has no further covariates added.

For native households, the estimate indicate an increase in moving propensity of around 0.3 percentage points if a native households gets a new non-west neighbor among their three nearest neighbors compared to receiving one slightly further away (ranks 4-6). This represents a 1.6 percent relative to the baseline moving rate of 20.23 percent moving rate. Importantly, this effect remains statistically significant and remarkably stable across different specifications. It is worth noting that this includes adding controls for real equivalised household income and/or real net wealth. These are two covariates I would expect to influence an incumbent household's propensity to move.

In contrast, Table 5 reveals that non-Western households show a substantially smaller response: 0.06-0.1 percentage points or around 0.5 percent relative to the baseline exit rate. This effect is not statistically significant for any specification and suggests that they are unaffected by the identity of a new different-type neighbor. This asymmetry in responses may reflect several factors. First, as Blair (2017) notes, outside options matter - if non-Western households perceive limited alternatives in equally desirable neighborhoods, they may be less responsive to receiving native neighbors. Second, the historical context of Danish immigration patterns means that non-Western households may already expect to live in mixed neighborhoods, whereas some native households hold different expectations.

Comparing these findings to Bayer et al. (2022), who examined racial preferences in the US, reveals interesting cross-country differences. While I find asymmetric responses where only native households exhibit significant responses, Bayer et al. (2022) find symmetric responses among both Black and White households, with magnitudes of 6 and 4 percent increases relative to their baseline exit rates within 2 years, respectively. My estimated effect for native households (1.6 percent) is substantially smaller than for either group. This difference may reflect institutional variation in housing market and integration policies, but also the distinct neighborhood contexts in which these "experiments" occur. As I have shown previously, most of these experiments in Denmark tend to happen in high-dense neighborhood, whereas Bayer et al. (2022) finds that these "experiments" mostly take place in suburban-sized neighborhoods. These contrasts suggest that while Schelling mechanisms operate across different settings, their magnitude and symmetry are shaped by local spatial, social and policy contexts.

Table 4: Estimates of Schelling behavior (native households)

	Move within 2 years (=100)					
	(1)	(2)	(3)	(4)	(5)	(6)
New diff neighbor $k_{nearest}$ v k_{near}	0.357***	0.368***	0.354***	0.328***	0.309***	0.320***
	(0.084)	(0.084)	(0.081)	(0.081)	(0.081)	(0.081)
N	5,365,811	5,365,811	5,365,811	5,365,811	5,365,811	5,365,811
Neighborhood-by-quarter FE	X	X	X	X	X	X
Mean of dependent variable	20.23	20.23	20.23	20.23	20.23	20.23
Number of neighborhoods	3444	3444	3444	3444	3444	3444
Income		X	X	X		X
Wealth					X	X
Tenure			X	X	X	X
Age				X	X	X

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors (in parenthesis) are clustered at the neighborhood level. The table reports the estimate of $\beta_1 - \beta_2$ from equation 6. Table A.3 contains the complete set of coefficients.

Table 5: Estimates of Schelling behavior (non-Western households)

	Move within 2 years (=100)					
	(1)	(2)	(3)	(4)	(5)	(6)
New diff neighbor $k_{nearest}$ v k_{near}	0.063	0.062	0.155	0.097	0.097	0.096
	(0.129)	(0.129)	(0.127)	(0.126)	(0.126)	(0.126)
N	1,795,109	1,795,109	1,795,109	1,795,109	1,795,109	1,795,109
Neighborhood-by-quarter FE	X	X	X	X	X	X
Mean of dependent variable	19.96	19.96	19.96	19.96	19.96	19.96
Number of neighborhoods	3332	3332	3332	3332	3332	3332
Income		X	X	X		X
Wealth					X	X
Tenure			X	X	X	X
Age				X	X	X

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors (in parenthesis) are clustered at the neighborhood level. The table reports the estimate of $\beta_1 - \beta_2$ from equation 6. Table A.4 contains the complete set of coefficients.

4.1 Heterogeneity

The full set of coefficients for tables 4 and 5, which can be found in tables A.3 and A.4, also indicate noteworthy heterogeneity, particularly along the income and wealth dimension. As mentioned previously, Blair (2017) emphasizes that outside options - which are often constrained by economic resources - also affects how households respond to neighborhood demographic changes. This motivates me to elicit insight into the differences in response by socioeconomic status (SES). I define SES as either "*low*" or "*high*":

1. **Low**: Households earn less than 200,000 DKK in real equivalised terms and was outside the labor marker *or* the best educated individual had less than or equal to 11 years of schooling.

2. **High**: Household earn at least 600,000 DKK in real equivalised terms and at least one member had full time work *or* the best educated individual had more than or equal to 18 years of schooling (corresponding to a completed long-cycle/MSc education.

The heterogeneity in response by SES is presented in table 6 and 7 for native and non-Western households, respectively. The first two columns in both tables denotes instances where the incumbent households fall within either of these categories. The next 4 columns "pairs" incumbent household by SES with incoming different-type neighbors that also fall within the categories above.

What is immediately clear from these tables is that the Schelling response is primarily driven by low-SES native households "paired" with low-SES non-Western household, who show an increase in moving propensity by around 0.56 percentage points when they receive a new non-Western neighbor or around 2.8 percent increase from their baseline exit rate. This effect is nearly twice the magnitude observed in the full sample. Furthermore, the remaining columns in table 6 also exhibit a fascinating pattern. It is incredibly rare for low-SES native households to receive high-SES non-Western and for low-SES non-Western households to receive high-SES native households, which corroborates the idea of the power of residential sorting that happens at the neighborhood level.

Table 6: Estimates of Schelling behavior (native households) by SES

	Move within 2 years (=100)					
	SES: Low	SES: High	SES: Low v Low	SES: Low v High	SES: High v High	SES: High v Low
	(1)	(2)	(3)	(4)	(5)	(6)
New diff neighbor $k_{nearest}$ v k_{near}	0.557***	0.287	0.558***	0.275	-1.050	0.036
	(0.090)	(0.297)	(0.105)	(0.453)	(0.887)	(0.442)
N	5,883,637	609,652	3,614,630	156,497	55,008	310,061
Neighborhood-by-quarter FE	X	X	X	X	X	X
Mean of dependent variable	18.89	20.75	17.83	17.01	23.65	20.22
Number of neighborhoods	3451	3450	3451	3248	2688	3446
Income		X			X	X
Tenure	X	X	X	X	X	X
Age	X	X	X	X	X	X

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors (in parenthesis) are clustered at the neighborhood level. The table reports the estimate of $\beta_1 - \beta_2$ from equation 6. Table A.5 contains the complete set of coefficients.

Table 7: Estimates of Schelling behavior (non-Western households) by SES

	Move within 2 years (=100)					
	SES: Low	SES: High	SES: Low v Low	SES: Low v High	SES: High v High	SES: High v Low
	(1)	(2)	(3)	(4)	(5)	(6)
New diff neighbor $k_{nearest}$ v k_{near}	0.008	-1.452*	0.020	0.491	-1.654	-1.551
	(0.115)	(0.624)	(0.126)	(0.552)	(1.381)	(0.803)
N	1,984,581	169,445	1,609,988	127,667	37,413	117,717
Neighborhood-by-quarter FE	X	X	X	X	X	X
Mean of dependent variable	18.33	25.58	18.26	18.12	28.08	26.20
Number of neighborhoods	3447	3135	3444	3361	2383	3087
Income		X			X	X
Tenure	X	X	X	X	X	X
Age	X	X	X	X	X	X

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors (in parenthesis) are clustered at the neighborhood level. The table reports the estimate of $\beta_1 - \beta_2$ from equation 6. Table A.6 contains the complete set of coefficients.

The spatial decay of effects shown in tables A.3 and A.4 provides additional support for the Schelling mechanism. The moving response decreases monotonically with the distance to new different-type neighbors, consistent with a model where the intensity of neighbor interactions diminishes with physical separation.

To assess the robustness of these findings, I follow Bayer et al. (2022) and estimate an alternative specification where I combine all control distance into a single category:

$$Y_{i,j,t} = \gamma_1 \mathbb{I}[e', k = n_{nearest}] + \gamma_2 \mathbb{I}[e', k = n_{control}] + \gamma Z_{i,j,t} + \omega_{j,t} + \epsilon_{i,j,t} \quad (12)$$

Where the coefficient of interest is $\gamma_1 - \gamma_2$, the difference in moving propensity following a new different-type households among their three nearest neighbor compared to rank 4 to 40. While this specification increases statistical power, as table 8 shows, my preferred specification distinguishes between neighbor ranks to better capture the spatial decay of effects that better aligns with the theory of Schelling (1971).

Table 8: Estimates of Schelling behavior, combined control

	Native	Non-Western
	Move within 2 years (=100)	
	(1)	(2)
New diff neighbor $k_{nearest}$ v $k_{control}$	0.981***	0.540***
	(0.070)	(0.123)
N	5,365,811	1,795,109
Neighborhood-by-quarter FE	X	X
Mean of dependent variable	20.23	19.96
Number of neighborhoods	3444	3332
Income	X	X
Tenure	X	X
Age	X	X

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors (in parenthesis) are clustered at the neighborhood level. The table reports the estimate of $\gamma_1 - \gamma_2$ from equation 12. Table A.7 contains the complete set of coefficients.

The evidence presented here provides empirical support for Schelling's theoretical prediction that even mild preferences regarding the identity of their neighbor - the *third* type of segregation as he called it - can generate segregation. The response of native households, in particular those with low-SES that tend to live in high-dense neighborhoods, indicate that residential sorting based on ethnicity remains an active mechanism in Danish housing markets.

5 Conclusion

This paper provides empirical evidence of Schelling behavior among Danish households by examining how the ethnicity of their nearest neighbors affects their propensity to move. Using a nearest neighbor research design by Bayer et al. (2022) that compares households within the same neighborhood to those who receive a new different-type neighbors "just down the road", I identify plausibly causal effects of Schelling behavior.

My results show that native Danish households increase their propensity to move within 2 years by approximately 0.3 percentage points, which translate into a 1.6 percent increase relative to their baseline exit rate. This effect is robust across specifications, including controls for household (equivalised) income and wealth in conjunction with tenure and age. In contrast, non-Western show a substantially smaller and statistically insignificant response to receiving native neighbors.

Importantly, there is significant heterogeneity embedded within these responses. I find that Schelling behavior is primarily concentrated among low-SES native household responding to other low-SES non-Western households. I find an effect of around 2.8 percent increase from their baseline exit rate, nearly twice the magnitude observed in the full sample. This suggests that residential sorting patterns are already pronounced in relatively "poorer" neighborhoods where resources and outside options are likely to be more constrained.

These findings contribute to our understanding of segregation dynamics in several ways. First, they provide causal evidence of Schelling's theoretical prediction that even mild preferences regarding the identity of your (nearest) neighbor can influence the choice of residence. Second, the socioeconomic gradient in responses highlight the intersection of ethnicity and economic resources.

My findings also suggest that the magnitude of Schelling behavior in Denmark is relatively modest compared to the US. Here, Bayer et al. (2022) finds symmetric responses in new different-type neighbors among Black and White American households with increases in moving propensities of 6 and 4 percent relative to their baseline exit rates within 2 years, respectively.

In conclusion, this research demonstrates individually motivated segregation, as theorized by Schelling over five decades ago, remains a relevant mechanism in shaping residential sorting patterns in Denmark today.

References

Bayer, Patrick et al. (2022). "Distinguishing Causes of Neighborhood Racial Change: A Nearest Neighbor Design". In: *Social Science Research Network*. DOI: 10.3386/w30487.

Bentley, Jon Louis (1975). "Multidimensional binary search trees used for associative searching". In: *Communications of the ACM* 18.9, pp. 509–517.

Bjerre-Nielsen, Andreas and Mikkel Høst Gandil (2024). "Attendance boundary policies and the limits to combating school segregation". In: *American Economic Journal: Economic Policy* 16.1, pp. 190–227.

Blair, Peter (2017). *Outside options (now) more important than race in explaining tipping points in US neighborhoods*. Tech. rep. University of Chicago.

Böhlmark, Anders and Alexander Willén (2020). "Tipping and the Effects of Segregation". In: *American Economic Journal: Applied Economics* 12.1, pp. 318–347.

Boje-Kovacs, Bence et al. (2024). *Immigrants and Native Flight: Geographic Extent and Heterogeneous Preferences*. Rockwool Foundation Research Unit.

Bøje-Kovács, Bence János et al. (2023). "Applying data for urban neighborhood development". In.

Caetano, Gregorio and Vikram Maheshri (2017). "School segregation and the identification of tipping behavior". In: *Journal of Public Economics* 148, pp. 115–135.

Card, David, Alexandre Mas, and Jesse Rothstein (2008). "Tipping and the Dynamics of Segregation". In: *The Quarterly Journal of Economics* 123.1, pp. 177–218.

Chetty, Raj, Nathaniel Hendren, and Lawrence F Katz (2016). "The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment". In: *American Economic Review* 106.4, pp. 855–902.

Damm, Anna P and Christian Dustmann (2014). "Does growing up in a high crime neighborhood affect youth criminal behavior?" In: *American Economic Review* 104.6, pp. 1806–1832.

Damm, Anna P and Marie L Schultz-Nielsen (2008). "Danish neighbourhoods: Construction and relevance for measurement of residential segregation". In: *Danish Journal of Economics (Nationaløkonomisk Tidsskrift)* 146.3, pp. 241–262.

Eechoud, Edward van (2024). *polars-grouper*. Version py-0.3.0. URL: https://github.com/Edwardvaneechoud/polars_grouper.

Grand, Carl le and Ryszard Szulkin (2002). "Permanent disadvantage or gradual integration: explaining the immigrant–native earnings gap in Sweden". In: *Labour* 16.1, pp. 37–64.

Hasager, Linea and Mia Jørgensen (2024). *Sick of Your Poor Neighborhood? Quasi-Experimental Evidence on Neighborhood Effects on Health.* Tech. rep. IZA Discussion Papers.

Mingarelli, Luca (2021). *Personal website — lucamingarelli.com.* (Accessed 26-09-2024). URL: https://lucamingarelli.com/Teaching/schelling.html.

Schelling, Thomas C (1971). "Dynamic models of segregation". In: *Journal of mathematical sociology* 1.2, pp. 143–186.

Statistics Denmark, DST (2022). *Definition af vestlige og ikke-vestlige lande.* URL: https://www.dst.dk/Site/Dst/SingleFiles/GetArchiveFile.aspx?fi=289262707644&fo=0&ext=befolkning (visited on 10/30/2024).

— (2024a). *DDU nomenclature.* URL: https://www.dst.dk/da/Statistik/dokumentation/nomenklaturer/ddu-udd (visited on 11/30/2024).

— (2024b). *Hvor i Danmark bor personer med udenlandsk oprindelse?* URL: https://www.dst.dk/analysepdf/45741 (visited on 01/30/2025).

— (2024c). *Statistikbanken (SOGN1).* URL: https://www.statistikbanken.dk/statbank5a/default.asp?w=2560g (visited on 10/30/2024).

— (2024d). *Times (FAMILIE_ID).* URL: https://www.dst.dk/da/statistik/dokumentation/times/moduldata-for-befolkning-og-valg/familie-id (visited on 10/30/2024).

Vink, Ritchie et al. (2025). *pola-rs/polars: Python Polars 1.25.2.* Version py-1.25.2. DOI: 10.5281/zenodo.15032415. URL: https://doi.org/10.5281/zenodo.15032415.

Wei, Ran, Sergio Rey, and Elijah Knaap (2021). "Efficient regionalization for spatially explicit neighborhood delineation". In: *International Journal of Geographical Information Science* 35.1, pp. 135–151.

Appendices

A Additional tables

Table A.1: Balance test (native)

	Income (1,000)	Net wealth (1,000)	Oldest HH member (years)	Tenure (days)	Employed	Educ. length (years)	HH size
New diff neighbor $k_{nearest}$ v k_{near}	-1.408***	-1.903***	-0.043*	1.936	-0.002**	-0.035*	-0.009***
	(0.254)	(0.402)	(0.020)	(5.999)	(0.001)	(0.014)	(0.002)
New diff neighbor k_{near} v k_{near} (omit.)							
New diff neighbor $k_{close,10}$ v k_{near}	-0.042	0.620	0.031	20.750***	0.001	0.010	0.005*
	(0.255)	(0.402)	(0.019)	(5.710)	(0.001)	(0.013)	(0.002)
New diff neighbor $k_{close,20}$ v k_{near}	0.498*	1.625***	0.109***	53.312***	0.002**	0.025*	0.015***
	(0.224)	(0.377)	(0.018)	(5.262)	(0.001)	(0.013)	(0.002)
New diff neighbor $k_{close,30}$ v k_{near}	1.429***	2.993***	0.228***	93.723***	0.005***	0.045**	0.029***
	(0.256)	(0.402)	(0.019)	(5.823)	(0.001)	(0.014)	(0.003)
New diff neighbor $k_{close,40}$ v k_{near}	1.805***	4.979***	0.343***	124.422***	0.004***	0.055***	0.040***
	(0.307)	(0.496)	(0.023)	(7.935)	(0.001)	(0.016)	(0.003)
N	5,365,811	5,365,811	5,365,811	5,365,811	5,365,811	5,365,811	5,365,811
Neighborhood-by-quarter FE	X	X	X	X	X	X	X
Mean of dependent variable	341.26	60.61	44.05	2646.87	0.84	10.11	1.70
Number of neighborhoods	3444	3444	3444	3444	3444	3444	3444

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors (in parenthesis) are clustered at the neighborhood level. The table reports the estimate of $\phi_k - \phi_2$ from equation 6 for each distance band k, for $k \in [1-3], [7-10], [11-20], [21-30], [31-40]$ from equation 11.

Table A.2: Balance test (non-Western)

	Income (1,000)	Net wealth (1,000)	Oldest HH member (years)	Tenure (days)	Employed	Educ. length (years)	HH size
New diff neighbor $k_{nearest}$ v k_{near}	-0.065	0.171	-0.061*	31.109***	0.001	-0.011	0.001
	(0.347)	(0.522)	(0.028)	(7.656)	(0.001)	(0.018)	(0.004)
New diff neighbor k_{near} v k_{near} (omit.)							
New diff neighbor $k_{close,10}$ v k_{near}	-0.075	0.514	0.064*	58.269***	0.002	0.024	0.026***
	(0.350)	(0.507)	(0.027)	(7.448)	(0.001)	(0.017)	(0.004)
New diff neighbor $k_{close,20}$ v k_{near}	-0.315	0.400	0.191***	111.810***	0.001	0.021	0.039***
	(0.319)	(0.467)	(0.026)	(6.707)	(0.001)	(0.015)	(0.004)
New diff neighbor $k_{close,30}$ v k_{near}	-0.760*	1.043	0.297***	157.738***		0.030	0.062***
	(0.348)	(0.536)	(0.028)	(7.760)	(0.001)	(0.018)	(0.005)
New diff neighbor $k_{close,40}$ v k_{near}	-1.064*	2.171**	0.430***	202.154***		0.013	0.064***
	(0.455)	(0.700)	(0.037)	(10.536)	(0.002)	(0.023)	(0.006)
N	1,795,109	1,795,109	1,795,109	1,795,109	1,795,109	1,795,109	1,795,109
Neighborhood-by-quarter FE	X	X	X	X	X	X	X
Mean of dependent variable	315.64	46.50	42.94	2298.84	0.84	12.11	2.11
Number of neighborhoods	3,332	3,332	3,332	3,332	3,332	3,332	3,332

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors (in parenthesis) are clustered at the neighborhood level. The table reports the estimate of $\phi_k - \phi_2$ from equation 6 for each distance band k, for $k \in [1-3], [7-10], [11-20], [21-30], [31-40]$ from equation 11.

Table A.3: Estimate of Schelling behavior (native households)

	Move within 2 years (=100)					
	<=25m					
	(1)	(2)	(3)	(4)	(5)	(6)
New diff neighbor $k_{nearest}$ v k_{near}	0.357***	0.368***	0.354***	0.328***	0.309***	0.320***
	(0.084)	(0.084)	(0.081)	(0.081)	(0.081)	(0.081)
New diff neighbor k_{near} v k_{near} (omit.)						
New diff neighbor $k_{close,10}$ v k_{near}	-0.259***	-0.258***	-0.174*	-0.169*	-0.170*	-0.168*
	(0.078)	(0.078)	(0.076)	(0.075)	(0.075)	(0.075)
New diff neighbor $k_{close,20}$ v k_{near}	-0.586***	-0.587***	-0.391***	-0.365***	-0.361***	-0.362***
	(0.072)	(0.072)	(0.069)	(0.068)	(0.068)	(0.068)
New diff neighbor $k_{close,30}$ v k_{near}	-1.341***	-1.348***	-1.056***	-0.984***	-0.972***	-0.979***
	(0.078)	(0.078)	(0.074)	(0.074)	(0.073)	(0.073)
New diff neighbor $k_{close,40}$ v k_{near}	-2.104***	-2.114***	-1.815***	-1.685***	-1.663***	-1.672***
	(0.096)	(0.096)	(0.088)	(0.086)	(0.086)	(0.086)
Income 200,000 DKK - 400,000 DKK (omit.)						
Income 400,001 DKK - 600,000 DKK		1.820***	1.777***	2.027***		2.227***
		(0.085)	(0.078)	(0.071)		(0.072)
Income 600,001 DKK - 800,000 DKK		4.142***	3.610***	4.145***		4.568***
		(0.197)	(0.198)	(0.189)		(0.189)
Income 800,001 DKK - 1,000,000 DKK		5.489***	4.547***	5.721***		6.227***
		(0.339)	(0.330)	(0.311)		(0.311)
Tenure < 1 year (omit.)						
Tenure $[1-2[$ years			-2.793***	-2.709***	-2.706***	-2.698***
			(0.106)	(0.105)	(0.105)	(0.105)
Tenure $[2-4[$ years			-6.405***	-6.095***	-6.051***	-6.037***
			(0.118)	(0.117)	(0.117)	(0.117)
Tenure $[4-6[$ years			-10.022***	-9.334***	-9.246***	-9.210***
			(0.127)	(0.127)	(0.126)	(0.126)
Tenure ≥ 6 years			-17.910***	-14.606***	-14.430***	-14.323***
			(0.125)	(0.124)	(0.123)	(0.123)
Oldest in HH, age 30-40 (omit.)						
Oldest in HH, age 41-50				-9.418***	-9.354***	-9.327***
				(0.087)	(0.089)	(0.088)
Oldest in HH, age 51-60				-12.896***	-12.595***	-12.661***
				(0.089)	(0.092)	(0.091)
Wealth -200,000 DKK - 0 DKK (omit.)						
Wealth 1 - 200,000 DKK					-1.654***	-1.694***
					(0.063)	(0.063)
Wealth 200,001 - 400,000 DKK					-1.389***	-1.856***
					(0.097)	(0.096)
Wealth 400,001 - 600,000 DKK					-1.836***	-2.508***
					(0.120)	(0.118)
Wealth 600,001 - 750,000 DKK					-2.080***	-2.889***
					(0.145)	(0.143)
N	5,365,811	5,365,811	5,365,811	5,365,811	5,365,811	5,365,811
Neighborhood-by-quarter FE	X	X	X	X	X	X
Mean of dependent variable	20.23	20.23	20.23	20.23	20.23	20.23
Number of neighborhoods	3444	3444	3444	3444	3444	3444
Income		X	X	X		X
Wealth					X	X
Tenure			X	X	X	X
Age				X	X	X

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors (in parenthesis) are clustered at the neighborhood level. The table reports the estimate of $\beta_k - \beta_2$ from equation 6 for each distance band k, for $k \in [1-3], [7-10], [11-20], [21-30], [31-40]$.

Table A.4: Estimate of Schelling behavior (non-Western households)

	Move within 2 years (=100)					
	<=25m					
	(1)	(2)	(3)	(4)	(5)	(6)
New diff neighbor $k_{nearest}$ v k_{near}	0.063	0.062	0.155	0.097	0.097	0.096
	(0.129)	(0.129)	(0.127)	(0.126)	(0.126)	(0.126)
New diff neighbor k_{near} v k_{near} (omit.)						
New diff neighbor $k_{close,10}$ v k_{near}	-0.381**	-0.384**	-0.154	-0.160	-0.159	-0.161
	(0.125)	(0.125)	(0.121)	(0.120)	(0.120)	(0.120)
New diff neighbor $k_{close,20}$ v k_{near}	-0.715***	-0.716***	-0.307**	-0.272*	-0.269*	-0.271*
	(0.115)	(0.115)	(0.112)	(0.111)	(0.111)	(0.111)
New diff neighbor $k_{close,30}$ v k_{near}	-1.294***	-1.292***	-0.735***	-0.674***	-0.673***	-0.673***
	(0.122)	(0.122)	(0.117)	(0.116)	(0.116)	(0.116)
New diff neighbor $k_{close,40}$ v k_{near}	-2.009***	-2.003***	-1.321***	-1.214***	-1.215***	-1.212***
	(0.143)	(0.143)	(0.135)	(0.131)	(0.131)	(0.131)
Income 200,000 DKK - 400,000 DKK (omit.)						
Income 400,001 DKK - 600,000 DKK		1.464***	1.268***	0.962***		1.054***
		(0.186)	(0.174)	(0.171)		(0.170)
Income 600,001 DKK - 800,000 DKK		3.955***	3.055***	2.434***		2.627***
		(0.446)	(0.423)	(0.417)		(0.416)
Income 800,001 DKK - 1,000,000 DKK		7.488***	6.284***	5.867***		6.124***
		(0.879)	(0.869)	(0.845)		(0.846)
Tenure < 1 year (omit.)						
Tenure $[1-2[$ years			-3.216***	-3.012***	-3.001***	-3.007***
			(0.189)	(0.187)	(0.187)	(0.187)
Tenure $[2-4[$ years			-7.436***	-6.769***	-6.759***	-6.753***
			(0.219)	(0.219)	(0.219)	(0.219)
Tenure $[4-6[$ years			-11.126***	-9.747***	-9.741***	-9.722***
			(0.241)	(0.240)	(0.240)	(0.240)
Tenure ≥ 6 years			-17.234***	-13.162***	-13.195***	-13.145***
			(0.227)	(0.226)	(0.226)	(0.226)
Oldest in HH, age 30-40 (omit.)						
Oldest in HH, age 41-50				-8.573***	-8.593***	-8.556***
				(0.145)	(0.145)	(0.145)
Oldest in HH, age 51-60				-12.999***	-13.045***	-12.999***
				(0.159)	(0.159)	(0.159)
Wealth -200,000 DKK - 0 DKK (omit.)						
Wealth 1 - 200,000 DKK					0.774***	0.770***
					(0.121)	(0.121)
Wealth 200,001 - 400,000 DKK					0.772***	0.464*
					(0.231)	(0.230)
Wealth 400,001 - 600,000 DKK					-0.506	-0.936**
					(0.293)	(0.289)
Wealth 600,001 - 750,000 DKK					-0.676	-1.191**
					(0.411)	(0.410)
N	1,795,109	1,795,109	1,795,109	1,795,109	1,795,109	1,795,109
Neighborhood-by-quarter FE	X	X	X	X	X	X
Mean of dependent variable	19.96	19.96	19.96	19.96	19.96	19.96
Number of neighborhoods	3332	3332	3332	3332	3332	3332
Income		X	X	X		X
Wealth					X	X
Tenure			X	X	X	X
Age				X	X	X

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors (in parenthesis) are clustered at the neighborhood level. The table reports the estimate of $\beta_k - \beta_2$ from equation 6 for each distance band k, for $k \in [1-3], [7-10], [11-20], [21-30], [31-40]$.

Table A.5: Estimates of Schelling behavior (native households) by SES

	Move within 2 years (=100)					
	SES: Low	SES: High	SES: Low v Low	SES: Low v High	SES: High v High	SES: High v Low
	(1)	(2)	(3)	(4)	(5)	(6)
New diff neighbor $k_{nearest}$ v k_{near}	0.557***	0.287	0.558***	0.275	-1.050	0.036
	(0.090)	(0.297)	(0.105)	(0.453)	(0.887)	(0.442)
New diff neighbor k_{near} v k_{near} (omit.)						
New diff neighbor $k_{close,10}$ v k_{near}	-0.124	-0.433	-0.169	-0.168	-0.657	-0.647
	(0.085)	(0.279)	(0.101)	(0.436)	(0.896)	(0.412)
New diff neighbor $k_{close,20}$ v k_{near}	-0.320***	-0.626*	-0.381***	-0.674	-1.247	-1.085**
	(0.076)	(0.255)	(0.091)	(0.394)	(0.793)	(0.373)
New diff neighbor $k_{close,30}$ v k_{near}	-0.959***	-1.240***	-0.886***	-1.005*	-2.234**	-1.671***
	(0.080)	(0.252)	(0.094)	(0.406)	(0.807)	(0.365)
New diff neighbor $k_{close,40}$ v k_{near}	-1.833***	-2.107***	-1.538***	-1.511***	-2.654***	-2.432***
	(0.088)	(0.272)	(0.103)	(0.413)	(0.785)	(0.405)
Tenure < 1 year (omit.)						
Tenure $[1-2[$ years	-4.380***	-0.771*	-3.968***	-3.885***	-0.078	-1.217**
	(0.107)	(0.306)	(0.133)	(0.572)	(0.931)	(0.441)
Tenure $[2-4[$ years	-8.764***	-1.776***	-8.377***	-8.913***	-1.169	-2.249***
	(0.124)	(0.304)	(0.148)	(0.534)	(1.064)	(0.449)
Tenure $[4-6[$ years	-12.323***	-2.975***	-11.889***	-13.019***	-1.531	-3.579***
	(0.136)	(0.317)	(0.161)	(0.557)	(1.152)	(0.472)
Tenure ≥ 6 years	-16.975***	-6.468***	-16.576***	-17.826***	-5.541***	-6.428***
	(0.138)	(0.295)	(0.159)	(0.468)	(0.946)	(0.443)
Oldest in HH, age 30-40 (omit.)						
Oldest in HH, age 41-50	-6.194***	-11.933***	-6.133***	-6.597***	-10.997***	-12.311***
	(0.069)	(0.265)	(0.081)	(0.294)	(0.935)	(0.311)
Oldest in HH, age 51-60	-8.884***	-15.402***	-8.646***	-8.521***	-15.411***	-15.981***
	(0.073)	(0.309)	(0.082)	(0.285)	(0.798)	(0.385)
Income 200,00 DKK - 400,000 DKK (omit.)						
Income 400,001 DKK - 600,000 DKK		0.282			-0.371	0.404
		(0.213)			(0.650)	(0.279)
Income 600,001 DKK - 800,000 DKK		2.468***			1.650*	2.649***
		(0.250)			(0.692)	(0.313)
Income 800,001 DKK - 1,000,000 DKK		4.055***			2.186*	4.125***
		(0.331)			(0.914)	(0.444)
N	5,883,637	609,652	3,614,630	156,497	55,008	310,061
Neighborhood-by-quarter FE	X	X	X	X	X	X
Mean of dependent variable	18.89	20.75	17.83	17.01	23.65	20.22
Number of neighborhoods	3451	3450	3451	3248	2688	3446
Income		X			X	X
Tenure	X	X	X	X	X	X
Age	X	X	X	X	X	X

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors (in parenthesis) are clustered at the neighborhood level. The table reports the estimate of $\beta_1 - \beta_2$ from equation 6 for each distance band k, for $k \in [1-3], [7-10], [11-20], [21-30], [31-40]$.

Table A.6: Estimates of Schelling behavior (non-Western households) by SES

	Move within 2 years (=100)					
	SES: Low	SES: High	SES: Low v Low	SES: Low v High	SES: High v High	SES: High v Low
	(1)	(2)	(3)	(4)	(5)	(6)
New diff neighbor $k_{nearest}$ v k_{near}	0.008	-1.452*	0.020	0.491	-1.654	-1.551
	(0.115)	(0.624)	(0.126)	(0.552)	(1.381)	(0.803)
New diff neighbor k_{near} v k_{near} (omit.)						
New diff neighbor $k_{close,10}$ v k_{near}	-0.256*	-0.239	-0.308*	0.166	-0.100	-0.133
	(0.109)	(0.593)	(0.120)	(0.512)	(1.296)	(0.808)
New diff neighbor $k_{close,20}$ v k_{near}	-0.603***	-0.526	-0.638***	-0.311	-0.690	-0.985
	(0.104)	(0.546)	(0.116)	(0.497)	(1.313)	(0.739)
New diff neighbor $k_{close,30}$ v k_{near}	-1.058***	-1.040	-0.953***	-0.617	0.018	-1.952**
	(0.114)	(0.558)	(0.123)	(0.501)	(1.225)	(0.736)
New diff neighbor $k_{close,40}$ v k_{near}	-1.715***	-2.439***	-1.591***	-0.699	-1.801	-2.669***
	(0.136)	(0.620)	(0.145)	(0.545)	(1.258)	(0.804)
Tenure < 1 year (omit.)						
Tenure $[1-2[$ years	-4.218***	-1.078	-4.069***	-5.014***	-0.804	-1.003
	(0.185)	(0.771)	(0.189)	(0.692)	(1.692)	(0.888)
Tenure $[2-4[$ years	-8.036***	-5.071***	-7.847***	-9.821***	-1.713	-5.929***
	(0.237)	(1.014)	(0.244)	(0.713)	(2.367)	(1.038)
Tenure $[4-6[$ years	-11.044***	-6.591***	-10.955***	-13.983***	-4.013	-7.442***
	(0.262)	(0.966)	(0.267)	(0.767)	(2.195)	(1.089)
Tenure ≥ 6 years	-14.118***	-10.743***	-13.979***	-17.203***	-10.131***	-11.544***
	(0.263)	(0.884)	(0.270)	(0.684)	(1.939)	(1.020)
Oldest in HH, age 30-40 (omit.)						
Oldest in HH, age 41-50	-5.342***	-10.033***	-5.372***	-5.737***	-8.446***	-11.070***
	(0.124)	(0.685)	(0.134)	(0.390)	(1.791)	(0.776)
Oldest in HH, age 51-60	-7.396***	-16.418***	-7.442***	-7.369***	-14.744***	-17.578***
	(0.147)	(0.779)	(0.158)	(0.429)	(1.476)	(0.946)
Income 400,001 DKK - 600,000 DKK		0.677			-1.214	0.441
		(0.689)			(1.268)	(0.863)
Income 600,001 DKK - 800,000 DKK		1.528*			1.658	1.285
		(0.670)			(1.279)	(0.808)
Income 800,001 DKK - 1,000,000 DKK		4.949***			3.702	5.874***
		(1.225)			(2.506)	(1.474)
N	1,984,581	169,445	1,609,988	127,667	37,413	117,717
Neighborhood-by-quarter FE	X	X	X	X	X	X
Mean of dependent variable	18.33	25.58	18.26	18.12	28.08	26.20
Number of neighborhoods	3447	3135	3444	3361	2383	3087
Income		X			X	X
Tenure	X	X	X	X	X	X
Age	X	X	X	X	X	X

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors (in parenthesis) are clustered at the neighborhood level. The table reports the estimate of $\beta_1 - \beta_2$ from equation 6 for each distance band k, for $k \in [1-3], [7-10], [11-20], [21-30], [31-40]$.

Table A.7: Estimates of Schelling behavior (native households), combined control

	Native	Non-Western
	Move within 2 years (=100)	
	(1)	(2)
New diff neighbor $k_{nearest}$ v $k_{control}$	0.981***	0.540***
	(0.070)	(0.123)
New diff neighbor $k_{control}$ v $k_{control}$ (omit.)		
Income 200,000 DKK - 400,000 DKK (omit.)		
Income 400,001 DKK - 600,000 DKK	2.020***	0.964***
	(0.071)	(0.171)
Income 600,001 DKK - 800,000 DKK	4.138***	2.438***
	(0.189)	(0.418)
Income 800,001 DKK - 1,000,000 DKK	5.711***	5.870***
	(0.312)	(0.845)
Tenure < 1 year (omit.)		
Tenure $[1-2[$ years	-2.791***	-3.077***
	(0.103)	(0.187)
Tenure $[2-4[$ years	-6.182***	-6.838***
	(0.116)	(0.216)
Tenure $[4-6[$ years	-9.426***	-9.819***
	(0.126)	(0.236)
Tenure ≥ 6 years	-14.707***	-13.241***
	(0.122)	(0.222)
Oldest in HH, age 30-40 (omit.)		
Oldest in HH, age 41-50	-9.427***	-8.579***
	(0.087)	(0.145)
Oldest in HH, age 51-60	-12.909***	-13.007***
	(0.089)	(0.159)
N	5,365,811	1,795,109
Neighborhood-by-quarter FE	X	X
Mean of dependent variable	20.23	19.96
Number of neighborhoods	3444	3332
Income	X	X
Tenure	X	X
Age	X	X

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors (in parenthesis) are clustered at the neighborhood level. The table reports the estimate of $\gamma_1 - \gamma_2$ from equation 12.

Table A.8: DDU nomenclature and education length in years

Education completed	Education length (years)
9th grade	10
Preparatory education (forberedende uddannelser)	11
Vocational education, base course (grundforløb)	11
Vocational education, second course (stud.komp.)	12
Vocational education, main course (hovedforløb)	14
High school	13
Short-cycle higher education	15
Medium-cycle higher education	16
Long-cycle higher education	18
Ph.D	21

Note: These denote the education length classification using DDU nomenclature from Statistics Denmark (2024a).

A.1 Distance to neighbors

It is reasonable to question why I choose to require exactly 25 meters or less for your nearest neighbor in the sample definition in section 3.6. To elicit discussion, consider the following estimation. I now expand my sample to include households who have their three nearest neighbor wihtin a distance 100 meters. I then include treatments in distance bands of 0-12.5, 12.5-25 and 25-100 meters. Thus, I am estimating:

$$\begin{aligned} Y_{i,j,t} = {} & \alpha_1 \mathbb{I}[e', k = n_{nearest_{0,12.5}}] + \alpha_2 \mathbb{I}[e', k = n_{nearest_{12.5,25}}] + \alpha_3 \mathbb{I}[e', k = n_{nearest_{25,100}}] \\ & + \alpha_4 \mathbb{I}[e', k = n_{control}] + \gamma Z_{i,j,t} + \omega_{j,t} + \epsilon_{i,j,t} \end{aligned} \tag{13}$$

Where the coefficients of interest are $\alpha_1 - \alpha_4$, $\alpha_2 - \alpha_4$ and $\alpha_3 - \alpha_4$. The "control" group are all other households who get new different-type neighbors from rank 4 to 40 in similar fashion to previous estimates. Table A.9 show the full set of coefficients of this model. Here you can clearly see the response to a new different-type neighbor mostly happens at a very close distance. To be "conservative", I decide to cap the maximum distance to the nearest neighbor at 25 meters to account for potential inaccuracies in the approach for identifying nearest neighbor described in section 3.3.

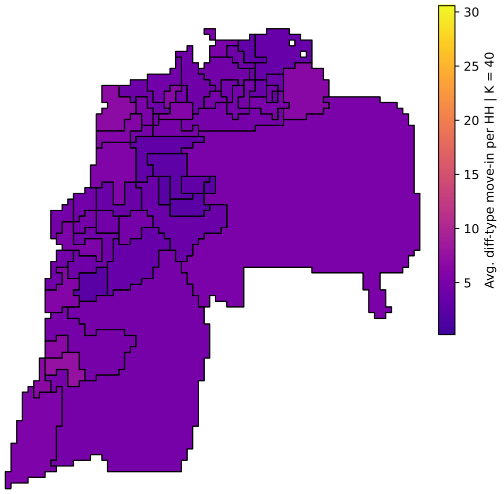
Table A.9: Estimates of Schelling behavior, distance to nearest neighbor $\leq$ 100m

	Move within 2 years (=100), $\leq$ 100m	
	Native	Non-Western
	(1)	(2)
New diff neighbor $k_{nearest,0}$ v $k_{control}$	1.702***	1.009***
	(0.078)	(0.122)
New diff neighbor $k_{nearest,25}$ v $k_{control}$	-0.817***	-1.101***
	(0.171)	(0.334)
New diff neighbor $k_{nearest,100}$ v $k_{control}$	-1.328***	-2.334***
	(0.228)	(0.535)
New diff neighbor $k_{control}$ v $k_{control}$ (omit.)		
Income 200,000 DKK - 400,000 DKK (omit.)		
Income 400,001 DKK - 600,000 DKK	1.445***	1.004***
	(0.060)	(0.152)
Income 600,001 DKK - 800,000 DKK	2.937***	2.333***
	(0.135)	(0.359)
Income 800,001 DKK - 1,000,000 DKK	3.900***	6.303***
	(0.229)	(0.756)
Tenure $<$ 1 year (omit.)		
Tenure $[1-2[$ years	-3.056***	-3.706***
	(0.095)	(0.173)
Tenure $[2-4[$ years	-6.495***	-7.833***
	(0.104)	(0.207)
Tenure $[4-6[$ years	-9.573***	-10.924***
	(0.113)	(0.226)
Tenure $\geq$ 6 years	-14.447***	-14.216***
	(0.112)	(0.217)
Oldest in HH, age 30-40 (omit.)		
Oldest in HH, age 41-50	-8.220***	-8.517***
	(0.087)	(0.135)
Oldest in HH, age 51-60	-11.068***	-12.439***
	(0.094)	(0.148)
N	6,543,928	2,036,442
Neighborhood-by-quarter FE	X	X
Mean of dependent variable	17.92	18.61
Number of neighborhoods	3451	3450
Income	X	X
Tenure	X	X
Age	X	X

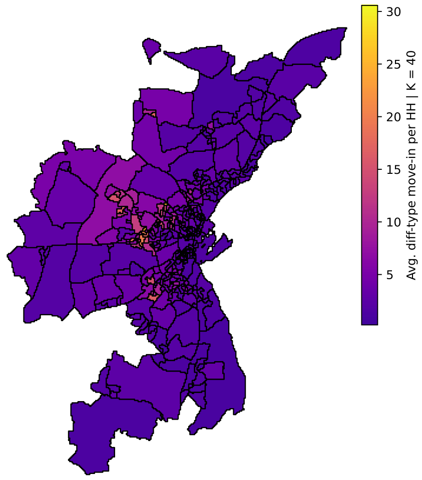
* p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors (in parenthesis) are clustered at the neighborhood level. The table reports the estimate of $\alpha_1 - \alpha_4$ from equation 13.

B Additional figures

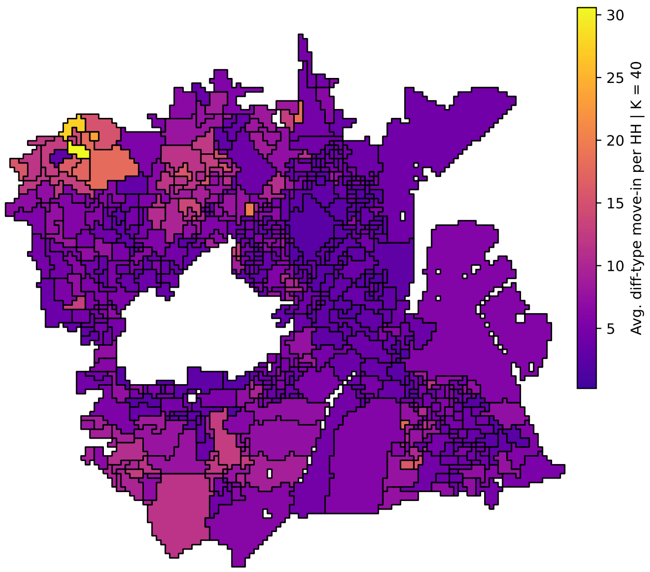
Figure B.1: Incidence of new different-type neighbors at the neighborhood level



(a) Ishøj



(b) Aarhus



(c) Copenhagen

Note: The figure show the variation in receiving a new-non West neighbor for native households at the *neighborhood* scale for three different municipalities. Neighborhoods not in the sample are greyed out. Household types are split up in three types (native/non-West/West), see section 1.1 for more details. Neighborhoods are defined in section 3.3.

C The Schelling Model

Consider a stylized example of Schelling's residential model: An $N \times N$ matrix $\mathbf{M}$ represents possible residential locations. Two types of agents b, r, graphically depicted as blue and red, respectively, are initially allocated a random location within this grid. For the purpose of this paper, these differ in terms of ethnicity.

The ratio between the two types of agents is constant $N_{b/r} = \frac{N_b}{N_r}$. Further, a constant share of residential locations are vacant, such that $N_{vacant} = N^2 - (N_b + N_r)$.

Each agent $i = b, r$ "optimizes" their residential location by how many of their K-nearest neighbors, i_K, are of the same type for some threshold τ. That is, for $s_K \leq \tau$, where a τ is a fixed global tolerance parameter for the maximum share of neighbors of a different type, each agent i will move until this condition is satisfied. In this very simple example, they move randomly to new vacant location at no cost.

Below is a parameterization of the model described above such that red is always in the minority[13]:

Table C.1: Schelling model parameters

Parameter	Value
N	200
$N_{\frac{b}{r}}$	1.25
N_{vacant}	0.1

From here, we can perform element-wise multiplication and summation for each agent in M to determine whether or not this satisfies the condition $s_i \leq \tau$. This process is called convolution and the operation can be done via the following kernel $\mathbf{K}$:

```
import numpy as np
KERNEL = np.array([[1,1,1,1,1],
                   [1,1,1,1,1],
                   [1,1,0,1,1],
                   [1,1,1,1,1],
                   [1,1,1,1,1]])
```

With the convolution operation defined as:

$$(K \star M)_{i,j} = \sum_{\substack{a \in \{-k_x, \ldots, k_x\} \\ b \in \{-k_y, \ldots, k_y\}}} K_{a+k_x, b+k_y} M_{i+a, j+b} \tag{14}$$

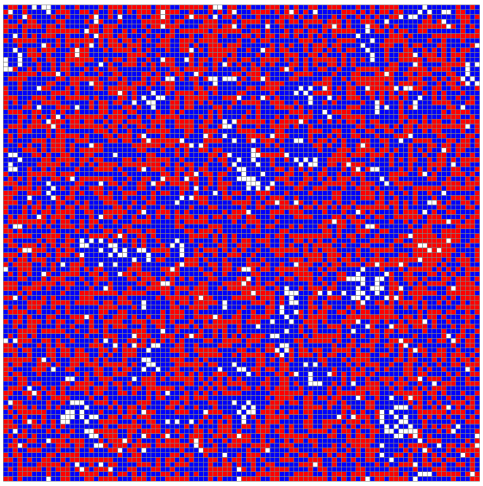
Where $k_x, k_y \in \mathbb{N}_0$.

[13] This implementation closely follows Mingarelli (2021), who very kindly made his code publicly available.

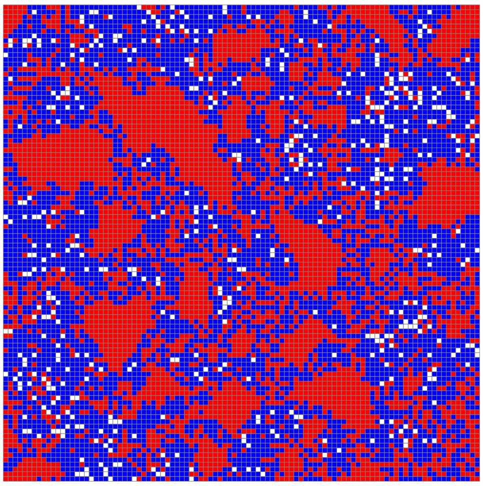
That is, I am considering the $k = 24$ nearest neighbors in reference to the tolerance parameter τ.

Figure C.1: Schelling model simulations by τ required same-type neighbors

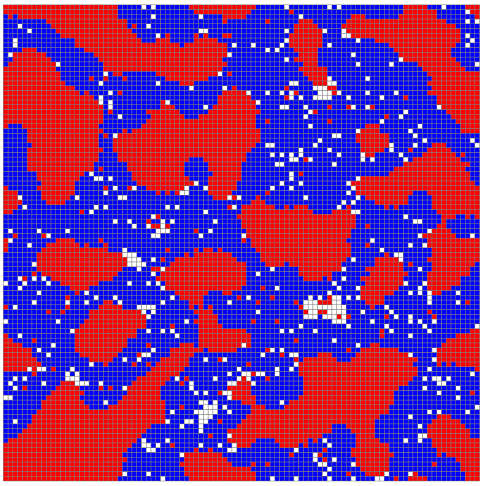
(a) "Assimilated" society, $\tau = 6$



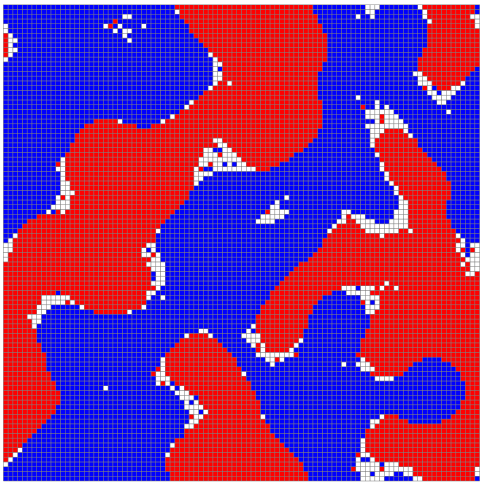
(b) "Integrated/segregated" society, $\tau = 8$



(c) "Segregated" society, $\tau = 10$



(d) "Gated" community, $\tau = 12$



Unsurprisingly, for sufficiently low levels of tolerances, one can observe fairly integrated societies. However, this pattern reverses once the tolerance parameter goes up, but not at levels one might expect.