Immigrants and Native Flight: Geographic Extent and Heterogeneous Preferences*

Bence Boje-Kovacs† Ismir Mulalic‡ Albert Saiz§
Vinicios Sant'Anna¶ Marie Louise Schultz-Nielsen‖

January 17, 2026

Abstract

Is immigrant segregation caused by native flight or self-isolation? If the former, which natives avoid immigrants and which types? Previous literature on ethnic segregation has focused on static models that characterize steady equilibria of local white shares. However, these models ignore churn and local sorting among majority populations. We address these issues using a dynamic model that contemplates constant mobility in and out of the neighborhood by both natives and minorities. Using a matched panel containing all individuals and properties in Denmark from 1987 to 2017, we exploit the quasi-random nature of refugee placements and simulated, exogenous Markov chain predictions to generate experimental variation in local immigrant arrivals. We find strong evidence of white flight, as measured by increased churn and net losses among the majority population, even at the building level. However, not everyone is upset by the presence of minorities. Flight is stronger among the elderly and in reaction to the arrival of low-income immigrants, regardless of their cultural background. As neighborhoods become densely populated with immigrants, young, low-income native citizens without children continue to move in disproportionately, but so do other immigrants.

Keywords: International Migration, Residential Segregation, Native Flight.
JEL codes: F22, J61, R21.

*The authors thank Statistics Denmark for the data. We thank seminar participants at the UMich, ROCKWOOL Foundation, the University of South Florida, Universitat Autònoma de Barcelona, Boise State, the University of Bern, the UEA North American and European Meetings, ARES, the 13th Annual Conference on "Immigration in OECD Countries", AREUEA National, and the Urban and Spatial Workshop for Junior Researchers for comments and suggestions. We gratefully acknowledge financial support from the ROCKWOOL Foundation. The usual disclaimer applies.

†Aalborg University, Department of the Built Environment, DK, email: bjbk@build.auu.dk.
‡Copenhagen Business School, Department of Economics, DK, email: imu.eco@cbs.dk.
§MIT, Center for Real Estate, Department of Urban Studies and Planning, USA, email: saiz@mit.edu.
¶California Polytechnic State University, Orfalea College of Business, USA, email: santanna@calpoly.edu.
‖ROCKWOOL Foundation Research, DK, email: mls@rff.dk.

1 Introduction

Over the past decades, many countries have experienced an uptick in immigration. According to the UNHCR, approximately 1 million immigrants reach Western Europe annually, and a similar number enter the USA. These inflows generate heated debates about their accommodation, assimilation, and impacts on destination economies. In fact, immigration has become *the* most salient political issue in developed countries. Anti-immigration sentiment, driven by a preference for ethnic homogeneity, has the potential to crystallize into conflict and the resurgence of immigrant ghettos (Cutler et al., 2008a). It is well-established that immigrants are likely to cluster locally, which may generate concerns about the appearance of "parallel societies" (Egge and Solhjell, 2018). Nevertheless, this tendency could be reinforced by *native flight* or *native avoidance* of areas with a higher concentration of immigrants (Betts and Fairlie, 2003).

In this paper, we use annual, individual, administrative data to uncover revealed preferences for homophily. We find causal evidence of native flight at both neighborhood and building levels. We also provide actionable policy targets by identifying the specific demographic patterns of this behavior. Ultimately, our results show that policies solely operating on immigrant "ghettos" may not fully succeed unless native behavior is also targeted.

Indeed, ethnic segregation raises serious policy questions, as it may result in the poor economic performance of minorities. Cutler and Glaeser (1997) shows that more segregated racial groups display worse economic outcomes compared to their counterparts in less segregated US cities. However, Edin et al. (2003) shows improved labor market outcomes in Sweden for immigrants living in segregated enclaves. A more recent body of literature has shown that higher concentrations of immigrants may delay the integration process (Cutler et al., 2008b). Nonetheless, the *quality* of information networks among minorities matters: having more highly-skilled ethnic neighbors improves employment prospects and wages (Damm, 2014).[1]

Previous research has explored the impact of immigrant arrivals on the mobility of natives *across cities*. Card (2001) shows that immigration inflows have only mild adverse effects on the earnings of low-skilled natives. Nonetheless, Borjas and Monras (2017) argue that this can be accounted for by labor market displacement: some native workers experience negative consequences and leave immigrant-dense metropolitan areas. Revisiting the widely-cited paper by Card (1990), Monras (2021) shows that native exits and

[1] Moreover, Borjas (1995) finds a positive relationship between the skill level of the first-generation immigrants in a neighborhood and the educational attainment as well as earnings of the second-generation. The study suggests that as immigrants' skill levels rise, their children's educational outcomes and income prospects improve accordingly.

reduced entry (native avoidance) fully account for the null effect of the Mariel boatlift on Miami's labor market. Amior (2020b), Amior (2020a), and Dustmann et al. (2017) also document native avoidance of immigrant cities in the US, the UK, and Germany, respectively. According to Derenoncourt (2022), individuals who grew up in commuting zones significantly impacted by the US African American Great Migration experienced a reduction in adult income with respect to individuals in other areas. Furthermore, Ortega and Verdugo (2022) show that native exits from immigrant-dense metro areas in France are not random: workers who lose out from competition by migrants are more likely to move out to other urban areas. In contrast, Foged and Peri (2016) finds no evidence of displacement of natives by immigrants *across municipalities* in Denmark, as low-skilled native workers switch to less manual-intensive occupations. Overall, cross-city studies point to the complex relationships between immigration and the native population as mediated by complementarities or substitutabilities in the metropolitan labor and housing markets (Saiz, 2003, 2007; Sá, 2015).

Recent work has expanded past research to include neighborhood outcomes, rather than focusing solely on metropolitan areas. This signifies a relative shift in interest away from labor markets and into social interactions. Specifically, changes in the ethnic composition of a neighborhood can be influenced by the preferences of majority-population residents (Miyao, 1979; Bond and Coulson, 1989; Han et al., 2022). This literature – heavily based on the US experience – draws from a strong tradition of research investigating the determinants of the urban segregation of the African American community.[2] Once the minority population in a neighborhood exceeds a certain *tipping point*, white flight may become more likely, leading to a significant local decline in the white population and complete segregation (Schelling, 1971; Bond and Coulson, 1989; Frankel and Pauzner, 2002; Card et al., 2008). However, racial segregation may also be partially attributed to the perception of lower-quality public services and urban amenities in such areas (Bayer et al., 2007a).[3]

As we argue throughout, the segregation literature has largely overlooked the dynamic nature of neighborhood change. Minorities are more likely to find vacant homes in neighborhoods with high turnover, which can lead to mistaking natural churn for white flight. While absolute segregation was the norm for Black people in mid-20th-century America, substantial periodic inflows and outflows of both minority and majority popu-

[2]Papers in this literature are too many to enumerate, but prominent examples include Taeuber and Taeuber (1965); Kain (1968); King and Mieszkowski (1973); Yinger (1976); Massey (1990); Cutler and Glaeser (1997); Cutler et al. (1999); Boustan (2010); Shertzer and Walsh (2019), and Derenoncourt (2022).

[3]Saiz and Wachter (2011) and Bayer et al. (2014) suggest that income differences are also an important determinant of segregation. This is also true for Denmark, where segregation by income level is significant (Gutierrez-i-Puigarnau et al., 2016; Mulalic and Rouwendal, 2020).

lations into the same neighborhoods are common nowadays. Furthermore, the types of people moving into and out of neighborhoods with large minority populations may differ greatly. This implies that white flight and gentrification could occur simultaneously in the same neighborhood if one group of the majority population flees while a larger group is attracted by the area's changing amenities.

Beyond black-white segregation, research about the drivers of neighborhood sorting by immigrants is also emerging. However, its depth and extent are still far from commensurate to the issue's social and political importance. For instance, Damm (2009) and Damm and Dustmann (2014) find that the density of co-nationals and availability of rental units drive immigrant location choices in Denmark. Focusing on net flows to and from urban neighborhoods, Saiz and Wachter (2011), Cascio and Lewis (2012), Accetturo et al. (2014), Moraga et al. (2019), Andersson et al. (2021), and Govind et al. (2025) show that immigrant inflows tend to be associated with net native outflows across neighborhoods in the US, Italy, Spain, Sweden, and Denmark. This effect is primarily observed in densely populated residential areas. In Sweden, Böhlmark and Willén (2020) finds evidence of tipping across neighborhoods. Moreover, the presence of low-income immigrants can be associated with perceptions of *vulnerability*, potentially stigmatizing a few neighborhoods (Andersson et al., 2023).[4] While this literature is very valuable, it arguably inherits many methods and limitations from classical studies of white-black segregation.

In this paper, we use comprehensive administrative data for the entire population of Denmark, providing micro-geographic information for over three decades, annually identifying all individuals, housing units, buildings, and neighborhoods. We test for the presence of native ethnic preferences and characterize their specific patterns. Studying native preferences for homophily and the segregation of minorities should not require much motivation due to their critical contemporary social and political importance. Nonetheless, we make a number of substantial empirical and methodological contributions to identify clear heterogeneous treatment effects that are ultimately important for policy.

First, we develop a novel dynamic theoretical framework combining individual preference shocks, moving costs, and heterogeneity in native distaste for the presence of immigrants. This framework illustrates the mechanisms through which immigration may alter a neighborhood's demographic composition and the patterns of local churn. It emphasizes the limitations of using aggregate, low-frequency ethnic shares to study native behavior and residential mobility in response to the arrival of minority populations. Our

[4]Note that there is a separate research literature on the impacts of neighborhood characteristics on immigrant outcomes – many of them in Northern European countries – (e.g. Edin et al. (2003); Åslund and Fredriksson (2009); Åslund et al. (2010, 2011); Damm (2014); White et al. (2016); Ansala et al. (2020), to which our paper does not contribute.

model reveals three separate patterns that characterize how natives respond to immigrant inflows: (i) replacement without displacement, (ii) flight and avoidance, and (iii) taste-based sorting.

Replacement without displacement occurs when incoming immigrants replace native residents who would naturally leave the neighborhood, due to the availability of housing vacancies. People move for many reasons related to life-cycle events, job relocation, or other external factors (Ahlfeldt et al., 2025). Naturally, neighborhoods with high churn are better able to host newly-arrived households.

Under native flight and avoidance, immigration triggers an accelerated demographic transition of residents beyond what would occur through natural turnover alone. This classic scenario arises when natives experience disutility from living in areas with a higher minority presence. In contrast, taste-based sorting considers nuanced heterogeneous preferences. Over time, those with stronger preferences for ethnic homogeneity will selectively relocate, while those who are more tolerant will tend to remain or move in. This triggers a sorting pattern among natives that would be undetected using aggregate measures of native resident shares.

Consequently, in our empirical exercise, we explicitly model the annual micro-individual probability of natives moving out of a neighborhood as a function of the lagged local share of immigrants.[5] Our approach avoids reverse causation from native exits to subsequent immigrant arrivals via increased vacancies. This problem plagues studies using decennial aggregate population changes. Modeling individual probabilities also allows us to control for a very rich set of household characteristics and local amenities that remained unobserved in previous studies – focusing on neighborhood aggregate dynamics – and which may confound the relationship of interest. For instance, minority-dense neighborhoods may have also been hosting natives with high propensities to churn.[6]

As a second contribution, we use alternative identification strategies that are highly consistent with causal interpretations. Endogeneity challenges plague the estimation of "white-flight" parameters (Moraga et al., 2019). On the one hand, contemporaneous shocks to a neighborhood may account for both changes in its immigrant density and the propensity of natives to move out. On the other hand, neighborhoods where immigrants used to cluster may have experienced different long-term trends, even without further

[5] To the best of our knowledge, except for Andersson et al. (2021), previous studies use *net* ethnic flows across neighborhoods.

[6] This problem is similar to that encountered in the study of gentrification because low-income households also have large residential turnover propensities to start with, so it is unclear if they were truly displaced. Pennington (2021) eloquently writes: *"Displacement happens to people; gentrification happens to places. Gentrification may happen without displacement (low-income incumbents willingly move and are replaced by higher-income newcomers), and displacement may happen without gentrification (push movers are replaced by newcomers from the same demographic)."*

immigration. Standard IV strategies that use shift-share predictions to estimate local immigrant arrivals address the former concern, but they may exacerbate the latter. To address these concerns, we provide estimates based on a novel Arrival-Stayer Markov Instrumental Variable (ASM-IV) approach. Our ASM-IV combines the random assignment of newly arriving refugees with mechanical mobility estimates. We also minimize concerns about over-reliance on a few initial shares, as raised by Bartik instruments (Goldsmith-Pinkham et al., 2020), by separating the variation arising from random arrivals. Additionally, our estimates control for neighborhood fixed effects and an exhaustive set of time-changing individual and neighborhood variables. They also control for trends that differ according to initial area characteristics.

Third, we examine various housing typologies because previous studies have revealed disparities in overall segregation across urban environments (Salazar, 2020). We document significant differences in behavior. Native flight is stronger between multifamily buildings located in denser areas. We precisely estimate zero effects within the single-family home subsample. While the distance between households in suburban areas may keep people apart (Putnam, 2000), it may also mitigate potential social tensions. Therefore, conventional local average treatment effects (LATE) obtained via shift-share instrumental variables (IVs) may overestimate white/native flight. This is because neighborhoods that used to host minorities tend to be in central, dense areas—precisely where the white/native out-migration response is largest.

Fourth, focusing on multifamily buildings, we find evidence of native flight across neighborhoods and *within* them at very local levels of residential interaction. Recent research in the US is turning to the study of the arrival of African American neighbors into adjacent houses (Bayer et al., 2025). However, as in many other parts of the world, the relevant micro-unit for social interaction in Northern and Eastern Europe is the housing complex: a collection of three- to four-story buildings with shared spaces that residents perceive as their living environment.[7] In addition to providing evidence on micro-interactions, these specifications let us control for neighborhood-and-year fixed effects, thereby dispelling most concerns about confounding unobserved local shocks.

Fifth, we combine our revealed-preference approach to individual mobility with non-linear analyses. Conventional tipping-point parameter estimates are based on net changes in the overall share of the majority population. Thus, they cannot identify true inflection points in exit rates among the existing white/native population whenever there are countervailing arrivals from white individuals with reduced distaste for minorities.

[7]Similar mid-rise, horizontally-expansive housing typologies are prevalent in some parts of the USA, such as courtyard apartments in Chicago (Bluestone, 2017) or garden apartments in California and the Southwest (Hise, 1995).

Sixth, our focus on individual behavior enables us to identify the marginal natives who flee from their immigrant neighbors. We observe strong evidence of heterogeneity; households led by older individuals tend to be more averse to foreigners. However, natives' education or income levels do not significantly influence their responses. Native residents of public housing also display a greater propensity to leave than homeowners or renters in private units. Furthering on the analysis of heterogeneous treatment effects, we also introduce local measures of the socio-economic and cultural distance between natives and the foreign-born. We find that results also vary based on immigrant characteristics: only low-income foreign-born arrivals spur native exit. In Denmark, cultural characteristics beyond income do not elicit native residential responses.

Seventh, we demonstrate that subsequent incoming residents in areas with high immigrant populations are more likely to be non-Western immigrants or young, low-income Danish citizens without children. To the best of our knowledge, this is the first analysis of majority-population arrivals to minority neighborhoods. Overall, our results suggest that policymakers should consider not only local ethnic composition but also sorting and changes within the native population in minority-dense areas.

Eighth, previous research has not been able to characterize the destinations of white leavers. As expected under the hypothesis of ethnic preferences, we confirm that natives who moved out of areas with high immigrant concentrations chose destinations with smaller immigrant shares.

Finally, we explicitly address concerns about multiple testing, a topic that is often overlooked in papers about segregation. When studying heterogeneous treatment effects and parameters for an increasing number of subpopulations, false positives are bound to arise randomly. However, we demonstrate that our results are robust to a split-sample strategy.

The remainder of the paper is organized as follows. Section 2 presents a dynamic framework of white flight. In Section 3, we introduce the data and provide descriptive statistics. Our estimation strategy is introduced in Section 4. Section 5 presents the main empirical results while Section 6 discusses heterogeneous treatment effects in native flight. In Section 7, we analyze how the share of non-Westerners influences the characteristics of incoming residents. Section 8 assesses the destination choices of people who moved out, while Section 9 concludes.

2 Churn, White Flight, and Sorting: Framework

In most conventional models of taste-based segregation, the majority population moves out of a neighborhood in response to the arrival of a minority population. These models tend to focus on comparative statics in equilibrium, where no one has further incentives to move. From its inception, this literature has focused on the stability of mixed-ethnicity neighborhoods (Miyao, 1978), providing insight into severe segregation and the existence of long-term, racially homogeneous neighborhoods. This line of work is also useful for rationalizing empirical patterns in decennial, neighborhood-level datasets that lack information on individual characteristics or dynamic residential decisions. However, they overlook critical factors contributing to ethnic segregation in contexts where residential mobility would be significant regardless of ethnic preferences.[8]

In this section, we develop a stylized dynamic model of residential mobility to examine how minority arrivals influence changes in neighborhood demographics. With utmost parsimony, it highlights the importance of using high-frequency individual-level data to disentangle the underlying mechanisms driving demographic shifts, such as replacement without displacement, taste-based native "flight," and ethnic-neutral churn. See Appendix A for proofs, complete derivations, and more detailed discussion.

2.1 Setup

Consider an economy with two locations: neighborhood A with fixed housing capacity at $\overline{H}^A$, and the rest of the world (or nation), with a perfectly elastic supply of dwellings. A continuum of residents of measure one choose their location each period based on their indirect utilities:

$$\begin{aligned} V_{i,t}^A &= \varepsilon_{i,t} - \Omega_t, \\ V_{i,t}^W &= 0 \end{aligned} \tag{1}$$

where $V_{i,t}^A$ is the indirect utility of individual i in period t derived from living in neighborhood A. $\varepsilon_{i,t} > 0$ is an idiosyncratic preference shock for A, drawn from an exponential distribution with parameter λ, and independent across residents and time. This annual shock captures changes in the proximity of each individual to their current job, or changes in tastes driven by lifetime events (such as family formation). Ω_t represents the endogenous cost of occupancy in A during period t, including rents and other non-monetary costs (such as waiting-list time for public housing). Residents in W obtain a reservation utility that is normalized at zero.

People incur a moving cost of $C \geq 0$ as they relocate from A to W.[9] Upon drawing

[8]Another family of models draws inspiration from choice models (McFadden, 2001) to estimate willingness to pay for living with people of other socio-demographic characteristics (Bayer et al., 2007b). We learn a lot about preferences from such empirical models but, again, they do not contemplate natural churn and its interactions with segregation in a multi-period, long-term dynamic setting.

[9]Introducing moving costs for relocation from W to A does not qualitatively affect the model's results

a new preference shock each year, individuals decide whether to stay or to move to the other region based on their current utilities. It is possible to consider an extension where people contemplate the option values derived from future taste shocks, but this would not add conceptual value as they can easily be subsumed by the moving cost, C.

Let $M_t^{W\to A}$ denote the mass of residents who move from W to A in period t, and let $M_t^{A\to W}$ capture those who move from A to W. Given the fixed housing capacity at A and the moving cost, a market-clearing equilibrium is characterized by an occupancy cost Ω_t^* such that inflows and outflows are equalized, ensuring that neighborhood A is fully occupied. In equilibrium, a constant mass $M_t^{W\to A} = M_t^{A\to W}$ of people move between the two locations every period, which we refer to as the *churn* rate of residents.

Proposition 1 ***(Churning and Moving Costs)*** *Higher moving cost C imply lower churn rates in and out of neighborhood A:*

$$\frac{\partial M^{A\to W}}{\partial C} = -\lambda \overline{H}^A e^{-\lambda(\Omega_t - C)} < 0. \tag{2}$$

Therefore, we have set up the environment to reflect the empirical facts that people of all groups routinely move in and out of their neighborhoods, but the degree of mobility may be heterogeneous across the latter.

Assume that this economy, initially in dynamic equilibrium,[10] receives an immigration shock of measure ϕ. The population now expands to $1+\phi$. Because housing supply in A is constrained, the arrival of more residents necessarily implies the expansion of the housing stock in the rest of the nation, W. We now use this framework to understand how changes in the presence of immigrants in A may affect native flight under different scenarios.

2.2 Neutral Native Ethnic Preferences

We first explore the scenario in which immigrants and natives exhibit the same distribution of preferences. Specifically, natives are indifferent to the presence of immigrants. We now express the *share* of residents in A who are immigrants in period t as $x_t^A = \phi_t^A / \overline{H}^A$, and $x_t^W = \phi_t^W / H^W$ as the average immigrant share in the rest of the country. Considering identical taste distributions, these measures will also determine the proportions of immigrants moving in and out of the neighborhood in the dynamic equilibrium. The

but mainly slows down the adjustment speed through reduced churn. For a more detailed discussion, see Appendix A.

[10] We use the terminology dynamic equilibrium to refer to an ergodic state. In such state, churn in and out of the neighborhood keeps happening. Nevertheless, while the share of immigrants in the neighborhood fluctuates period-by-period, it tends to converge to the same average over time.

evolution of immigrant presence in A can be written as:

$$\Delta\phi_t^A = M^{W\to A} \cdot x_{t-1}^W - M^{A\to W} \cdot x_{t-1}^A, \tag{3}$$

where $\Delta\phi_t^A$ is the change in the measure of immigrants living in A at time t. The first term captures immigrants moving into A (arrivals) while the second term represents immigrants leaving A (leavers).

Proposition 2 ***(Dynamic Equilibrium in the Share of Immigrants)*** *In the long run, the immigrant share in A converges to an equilibrium share, x^*, which is equivalent to the proportion of immigrants in the population, $\frac{\phi}{1+\phi}$:*

$$\Delta\phi_t^A = 0 \iff x_{t-1}^W = x_{t-1}^A \iff x^* = \frac{\phi}{1+\phi}. \tag{4}$$

Because immigrants and natives have the same preferences, residential sorting between locations happens only because of shocks to jobs or lifestyle amenities. Regardless of the initial distribution of immigrants in A, their presence will converge to their long run national average. In the short to medium term, however, this scenario generates a pattern of replacement without displacement. If, for example, all immigrants were initially allocated to W, the immigrant share in A will rise commensurate with voluntary native exits. This mechanism generates increasing immigrant concentrations in A without direct displacement. However, an observer relying on aggregate data might mistakenly attribute this pattern to "native flight," specially in neighborhoods with high natural churn.

Proposition 3 ***(Baseline Residential Churn and Speed of Demographic Transition)*** *Neighborhoods with low baseline moving costs, and consequently high churn (Proposition 1), experience faster demographic transitions to the dynamic equilibrium. Moreover, they systematically exhibit higher immigrant shares throughout the transition period.*

This result implies that, even when they have preferences identical to those of natives, immigrants initially cluster in the neighborhoods characterized by high residential turnover. Only over time does the distribution of immigrants become evenly distributed across space. Variation in underlying churn rates can lead to misleading conclusions about excessive local segregation and preference-based sorting during the transitional period.

2.3 Native Tastes for Homophily

We now assume homophilic residential ethnic tastes among natives. For simplicity, immigrants maintain the same preferences as before. However natives experience a utility *penalty* that is a function of the fraction of immigrants in neighborhood A. The indirect utility function for natives (N) in period t becomes:

$$\begin{aligned} V_{i,t}^{A}(N) &= \varepsilon_{i,t} - \Omega_t - \delta \cdot ln(1 + x_t^A), \\ V_{i,t}^{W}(N) &= 0, \quad \forall i \end{aligned} \tag{5}$$

where $\delta > 0$ represents a parameter capturing natives' distaste for living in A as the share of immigrants there (x_t^A) grows.[11]

Proposition 4 ***(Distaste and Native Avoidance)*** *The larger the distaste parameter δ or the share of immigrants x^A, the fewer natives will move to A, denoted by $M_t^{W\to A}(N)$:*

$$\frac{\partial M_t^{W\to A}(N)}{\partial \delta} < 0 \qquad \& \qquad \frac{\partial M_t^{W\to A}(N)}{\partial x^A} < 0. \tag{6}$$

Proposition 5 ***(Distaste and Native Flight)*** *The larger the distaste parameter δ or the share of immigrants x^A, the larger the measure of natives that will move out of A, denoted by $M_t^{A\to W}(N)$:*

$$\frac{\partial M_t^{A\to W}(N)}{\partial \delta} > 0 \qquad \& \qquad \frac{\partial M_t^{A\to W}(N)}{\partial x^A} > 0. \tag{7}$$

As shown in the online Appendix— the long-term dynamic ethnic equilibrium is such that inflows and outflows by ethnic group are equalized. This yields an equation that *implicitly* defines the long run immigrant share in neighborhood A (x^*) as a function of the model's parameters:

$$x^*\left(1 + (1 - x^*)\overline{H}^A\left(e^{\lambda C} - 1\right)\right) = (1 - x^*)(1 + x^*)^{\lambda\delta}\left(\phi + x^*\overline{H}^A\left(e^{\lambda C} - 1\right)\right) \tag{8}$$

We are now in a position to state:

Proposition 6 ***(Taste-Based Segregation)*** *In a long-run dynamic equilibrium the share of immigrants in A (x^*) is higher if natives have a distaste for living with immigrants.*

$$x^*_{\delta=0} < x^*_{\delta>0}. \tag{9}$$

This scenario parsimoniously captures "classical" white flight dynamics.[12] However, the dynamic specification helps us distinguish the nature of the long-term equilibrium from

[11]Replacing the log penalty in Equation (5) with a linear or exponential specification yields similar qualitative results but introduces undesired notational complexity. For a more detailed discussion, see Appendix A.

[12]Our model rules out a fully segregated neighborhood ($x^A = 0$ or $x^A = 1$). This is a direct consequence of our setting in which residents randomly draw ε's from an exponential distribution. Even for large

the speed at which we reach it. Specifically, per Proposition 1, convergence to the equilibrium will be slow in areas with high mobility costs. Therefore, rapid ethnic demographic transitions after reaching potential tipping points (Card et al., 2008) are a sufficient, but not necessary, condition for the existence of the latter.

In most countries outside of the U.S., household mobility is much weaker. This implies that larger shocks are needed to prompt natives to leave their homes and neighborhoods, even if their perceived quality of life decreases due to the arrival of immigrants. However, as they repeatedly keep sampling from the neighborhood-matching distribution, a combination of mobility shocks and homophilic preferences will eventually prompt them to move out. In such environments, tipping points may have been crossed, and some neighborhoods may be heading inexorably toward a maximal segregation equilibrium. Nevertheless, this could take a very long time. Other random economic fluctuations will inevitably affect these neighborhoods in the coming decades, so a fully segregated equilibrium may never be reached. In such cases, researchers will only observe partial white flight.

2.4 Heterogeneous Native Preferences

In the presence of native taste heterogeneity, analyses that focus exclusively on the share of minority and majority populations become less relevant. To see this, consider the existence of two types of natives: *Segregationists* and *Neutrals*. Segregationists dislike living with immigrants, as in Equation (5). Neutrals are indifferent to the presence of immigrants, and therefore have the same preferences as them, captured by Equation (1). Denote the proportion of segregationists in the native population as θ.

Proposition 7 ***(Taste-Based dynamic Sorting)*** *In a scenario where natives have different preferences regarding immigrants, native flight is partially offset by neutral native inflows. This implies that the dynamic share of immigrants in A, x^*, will be higher than in the scenario where all natives are indifferent to immigrants* ($\theta = 0$) *but lower than in the scenario where all natives dislike immigrants* ($\theta = 1$)*:*

$$x^*_{\theta=0} < x^*_{0<\theta<1} < x^*_{\theta=1}. \tag{10}$$

The most notable aspect of the demographic transition in this scenario is the replacement of one type of native by another. However, as shown in the appendix, there are a variety of possible paths for the speed of sorting and ethnic transition in neighborhoods. These

immigrant "distaste" parameters, a small measure of natives will inevitably still draw large offsetting ε's. However, this measure may become arbitrarily close to zero as δ grows, and hence this parsimonious framework also encompasses full segregation as an extreme case.

paths depend on the distribution of local productivity/amenity shocks, mobility costs, the intensity of homophilic preferences, and the proportion of tolerant natives in the country.

2.5 Beyond Aggregate Minority Shares

Our dynamic framework highlights the limitations of using low-frequency, aggregate neighborhood data—such as decennial censuses—to understand the impact of minority arrivals on residential sorting. While such data may reveal net changes in native and immigrant populations, they are limited in their ability to distinguish between the various mechanisms at work. These mechanisms include whether the decline in native presence reflects pre-existing high churn (replacement without displacement), job or amenity shocks, native flight, taste-based sorting, or a combination of these phenomena. In addition, this type of data makes it easier to confuse the nature of the long-term equilibrium with the speed at which the neighborhood will reach it.

Most problematically, the offsetting flows predicted by heterogeneous preferences remain invisible in statistics about local net ethnic shares. This can lead researchers to underestimate the extent of native flight and anti-immigrant sentiment in the population. The typologies of natives who are simultaneously fleeing and moving into a neighborhood may differ greatly, which could be as important as the arrival of immigrants to determining the neighborhood's future success. Furthermore, ethnic preferences may depend on the characteristics of immigrants, which are not fully observable in aggregate data.

In turn, failing to account for replacement without displacement could also lead researchers to overestimate flight responses of the native population, especially in areas where mobility was already high. We therefore argue that, to properly understand the dynamics of ethnic segregation, it is necessary to use high-frequency, micro-level panel data that tracks individuals or households at high frequencies, as we do below.

3 Data

3.1 Statistics Denmark: a Population Registry

We exploit the richness of annual longitudinal administrative data from 1987 to 2017 for the register of the entire population processed by Statistics Denmark (DST). DST uses individual IDs to collate data from the countries' numerous administrative registries, including residential, income tax, cadastre, migration, social security, judicial, health, employment, and educational data. For this study, we can link a host of individual demographic variables – including age, gender, education, marital status, number of children, income, and migration status – to residence identifiers, i.e., housing unit, building, and neighborhood where each individual lives.[13] We also observe housing tenure types and attributes such as building age, number of rooms in the unit, number of units in a building, and property type (public versus private and single- versus multifamily home). Our sample thus consists of a matched individual-and-property panel data set for the period 1987–2017, covering the universe of all individuals and properties in Denmark.

3.2 Neighborhoods

We define two types of neighborhoods by size according to the georeferenced clustering procedure developed by Damm and Schultz-Nielsen (2008). Neighborhoods are delineated by existing physical barriers and constructed to be homogeneous by the number of residents and housing typology. There are 9,401 small neighborhoods consistently defined over time and populated by at least 150 households, with an average area of 677 m^2. Encompassing several adjacent small areas, 2,296 large neighborhoods are defined to contain at least 600 households. We adopt the latter definition, as it is less sensitive to outliers but granular enough in urban areas, as can be seen for Copenhagen in Figure 1.[14] Nonetheless, Table D.1 in Appendix D shows that our results are robust to using the alternative small-neighborhood definition.

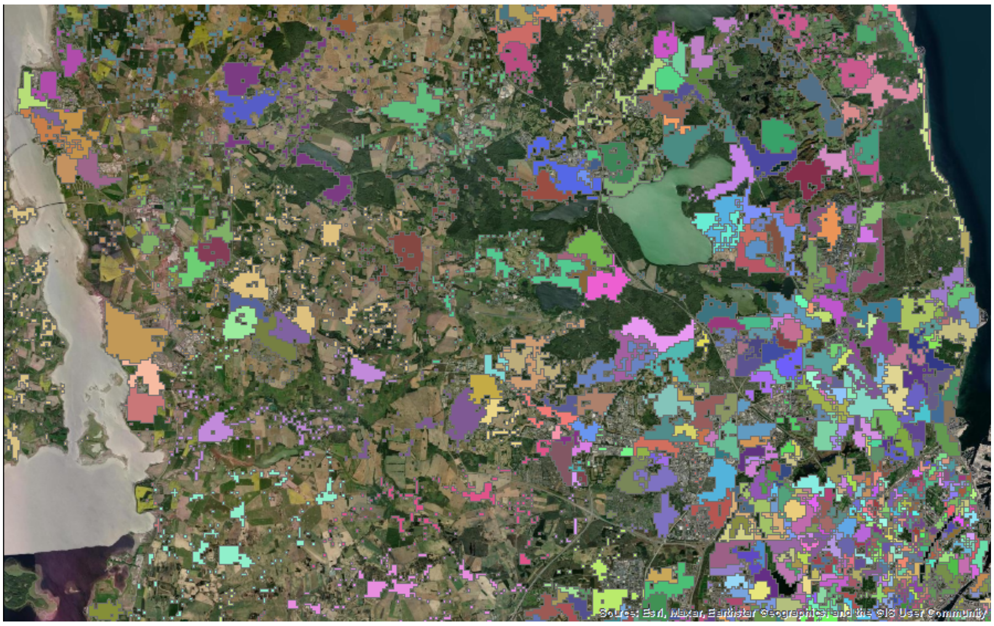


Figure 1. Large Neighborhoods in Central Copenhagen based on the definitions of large neighborhoods as constructed by Damm and Schultz-Nielsen (2008).

[13]To ensure confidentiality, Statistics Denmark does not provide the exact residence addresses but instead replaces them with scrambled identifiers.

[14]Appendix B shows the spatial distribution of centroids for both large (Figure B.2) and small (Figure B.3) neighborhoods in the whole country.

3.3 Definitions and Descriptive Statistics

Per the DST, immigrants are defined as individuals born outside of Denmark and with neither parent who is a Danish citizen born in Denmark. We also observe citizenship and country of origin. According to DST's official definitions, Western immigrants come from the EU/EEA, European microstates (Andorra, Liechtenstein, Monaco, San Marino, and the Vatican), Switzerland, the UK, Canada, the USA, and Australia. Non-Western immigrants hail from any other country. To study native household behavior, we focus on adult non-immigrant heads of households. Due to mortality and new household formation, our final dataset consists of an unbalanced panel of over 52 million observations on more than 4.8 million unique heads of household.

The data provide an ID for each *residential structure* in the country. They also identify each *street door* and apartment's *entry door* by building. We categorize single-family detached buildings as structures with a unique *entry door* – which coincides with the *street door*. Multifamily structures, on the other hand, include multiple *entry doors*. Most multifamily units are located in building complexes, which are sets of adjacent buildings within the same residential structure. The tall buildings or modernist complexes that characterize housing apartments/condos in the US, Southern Europe, Latin America, and Asia are less prevalent in many central and northern European countries. They are very rare in Denmark. Figure 2 illustrates the nature of Danish building complexes: collections of identical, 3-6 story-high buildings attached horizontally. The figure – extracted from Google Maps© – is one such ensemble with 5 street doors, each facilitating access to six separate housing units (totaling 30 apartments). Local market participants recognize each structure as a distinctive "place" or "address."

Figure 2. Building Complex Typology in Denmark. This figure shows an example of a residential building complex in the Copenhagen Metropolitan Area drawn from Google Maps©.

In most specifications, we register a *move-out* whenever a head of household changes residence from one year to the next, *and the new address is located in a different neigh-*

borhood. Conservatively, moves within a neighborhood are not classified as native flight. Nevertheless, we compute three alternative dummies corresponding to (1) moving out of an address, (2) a building, and (3) a neighborhood (shown in the first three rows of Table 1).

11% of household heads changed address each year, often moving to a different neighborhood (8%). The remainder of Panel A in Table 1 summarizes the data for household heads: they were 48 years old on average, and 59% were male. Over one-third held vocational education, while 24% achieved higher educational levels. Interestingly, only 3.7% were unemployed, and 1% were students, with 29% out of the labor force. More than one-fifth were either divorced or widowers, and 45% were cohabiting. On average, there were 1.6 adult individuals per household and 0.5 children. A majority of household heads owned their homes and stayed at the same address for more than six years.

Table 1 also shows that housing units in Denmark are 53 years old on average, and typically contain less than four rooms. One-third of households live in large building complexes (with more than ten units), and approximately 19% live in public housing. The average distance to a city center is about 5 km, ensuring good access to urban amenities (e.g., banks, hospitals, primary schools, grocery stores, and restaurants).[15]

It is crucial to have good and context-relevant metrics of native exposure to immigrant residents. Our rich data allows us to create a precise annual measure: the share of non-Western immigrants – which we henceforth refer to as "immigrants" – in the previous year by neighborhood or building complex.

Table 2 shows exposure statistics for all neighborhoods and building complexes. The mean share of immigrants by neighborhood was 4.7%, but the standard deviation is relatively high. Some areas were mostly populated by immigrants (the maximum is 80%). The figures at the building level reveal an unweighted mean share of immigrants of only 3.2%, again with a high standard deviation. This implies that non-Western immigrants disproportionately live in multi-unit buildings.[16]

Figure 3 shows the evolution of the share of immigrants and the share of natives changing residences between 1988 and 2017. The percentage of immigrants rose from 2% to more than 10% in 2017, with the share of non-Western immigrants increasing almost linearly. By contrast, the probability of native mobility was stationary, if slightly pro-cyclical.

Finally, we calculated the correlation between the proportion of non-Western immigrants and the native turnover rate by neighborhood, which is 0.10% (with a standard

[15]See Appendix C for a description of how we calculated neighborhood amenities.

[16]Focusing only on data from buildings with more than 10 units, the average share of immigrants by neighborhood increases to 5.8%, while the average share by building complex becomes 6.1%.

Table 1. Descriptive Statistics

Variable	Mean	SD	Median
Panel A - Household Heads			
Moved	0.112	0.316	0
Moved building	0.111	0.314	0
Moved neighborhood	0.080	0.271	0
Age	48.321	16.375	47
Female	0.407	0.491	0
Married or stable relationship	0.453	0.498	0
Divorced, dissolved partnership, or widow	0.228	0.420	0
Higher education	0.244	0.429	0
Vocational education	0.370	0.483	0
Out of the labor force	0.286	0.452	0
Student	0.011	0.103	0
Unemployed	0.037	0.188	0
Household disposable income (log)	12.526	0.590	12.548
Number of children in the household	0.501	0.871	0
Number of residents in the household	1.632	0.613	2
Renter	0.444	0.497	0
Length of occupancy (years)	6.751	6.026	5
Panel B - Properties			
Number of rooms (within unit)	3.814	1.61	4
Number of units (within property)	43.727	105.221	1
Large building complex (10 or more units)	0.368	0.482	0
Public house	0.188	0.391	0
Property age	53.26	37.707	43
Panel C - Neighborhoods			
Number of banks	1.912	3.934	0
Number of hospitals	0.218	0.673	0
Number of primary schools	1.425	2.045	1
Number of grocery stores	2.820	5.024	1
Number of restaurants	13.032	33.95	2
Number of doctors	9.483	19.524	1
Number of entertainment business	6.104	14.488	2
Share of employed residents	0.604	0.094	0.61
Proximity to city center (km)	4.973	5.757	2.450
Proximity to roads (km)	15.400	17.691	11.913
Proximity to the coast (km)	6.369	8.252	3.192
Proximity to waste (km)	26.949	23.230	17.881
Proximity to lake (km)	0.409	0.317	0.343
Proximity to forest (km)	0.326	0.335	0.238
Proximity to train station (km)	4.527	12.796	1.417
Panel D - Immigrant Presence			
Share of non-Western immigrants in neighborhood	0.044	0.062	0.023
Share of non-Western immigrants in building complex	0.032	0.081	0
Predicted share of refugees in neighborhood	0.012	0.02	0.006

Note: Number of observations is 52,025,214 (4,168,843 heads of household).

error of 0.024%). This positive correlation suggests that researchers must proceed with caution in this type of study because non-Western immigrants may have moved dispro-

Table 2. Descriptive Statistics for Immigrant Presence

	N	Mean	SD	Median
Panel A - Complete Sample				
Neighborhood				
Number of residents	68,850	1,624	628	1,488
Immigrant share	68,850	0.047	0.071	0.023
Building				
Number of residents	31,259,382	3.346	15.462	2
Immigrant share	31,259,382	0.003	0.033	0
Panel B - Large Building or Complex (10 or more units)				
Neighborhood				
Number of residents	52,094	1,601	639	1450
Immigrant share	52,094	0.058	0.078	0.031
Building				
Number of residents	494,212	60.063	98.745	31
Immigrant share	494,212	0.061	0.101	0.021

portionately to areas that would have experienced high residential turnover, regardless of immigration.

3.4 Other Data Sources

Our analysis deploys additional data sources to categorize the countries of origin of non-Western migrants. Our goal is to incorporate variables that could play a role in the assimilation of immigrants to the Danish culture and in explaining native responses.

Time Varying Amenities To define consumer amenities, we utilize Danish administrative register data containing the entire universe of firms from 1987 to 2017 (see Appendix C for details). We follow the evolution of three types of retail establishments by small neighborhood and year: food services, entertainment, and grocery stores. We also condition for the number of primary schools, doctors, and banks in the neighborhood.

By conditioning on contemporaneous consumption amenities, we control for changes in local livability that may affect foreign and indigenous population flows differentially. Conservatively, we also aim to mute any indirect effects going through endogenous changes in local amenities (Waldfogel, 2008). However, these efforts turned out not to be relevant in practice, as results are unchanged when we omit these variables.

Time Invariant Amenities In order to capture first-order geographic attributes, whose valuations may be changing during this period, we interact time dummies with distances from the small neighborhood's centroid to: each city/town center, highways, the ocean, waste facilities, a lake, forests, and train/subway stations.

Origin Country's Religious Composition We use information from the religious

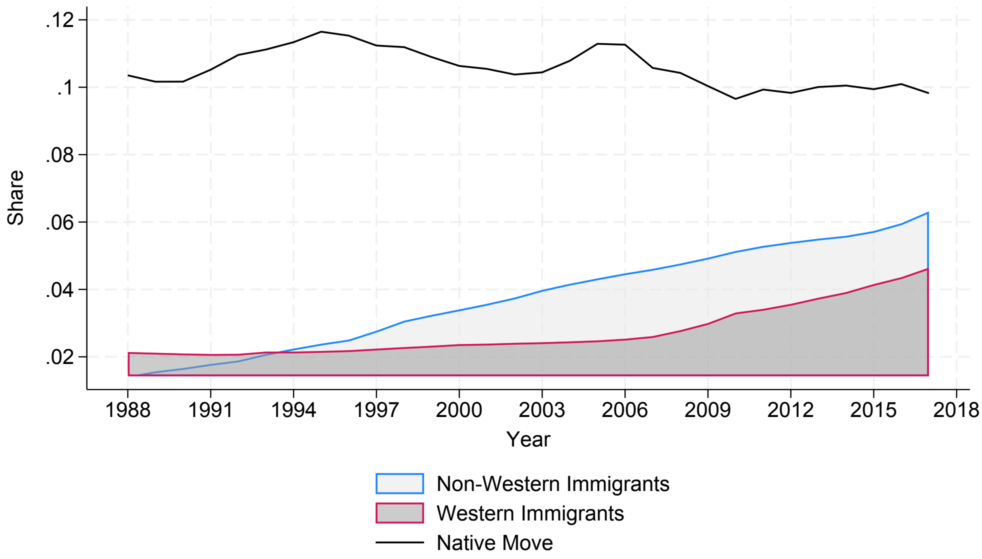


Figure 3. Long-Term Trends in Immigrant Share and native relocation, 1988-2017. This figure plots the average share of Danish residents moving to a different address each year (black line) and the shares of Western (dark-gray) and non-Western (light-gray) immigrants in the country's population.

composition of countries from Pew Research Center (2015). We classify countries as "Muslim" if more than 50% of their population is Muslim.

Language Proximity We use an index of language proximity between Danish and the origin countries from *Centre D'Estudes Prospectives et d'Informations Internationales* (CEPII), available from Melitz and Toubal (2014). We use the lexical similarity between 40 words compiled by the Automated Similarity Judgment Program (ASJP). According to CEPII, this measure is better suited to compare languages in different families.

Facebook Connectedness Index This is a measure of international social connectedness between Denmark and other countries, obtained from Bailey et al. (2021). These authors use de-identified administrative data from social media (Facebook) to measure the relative probability of a friendship link between Facebook users by country dyad.

Country Income Level We use the World Bank's country classifications into four income groups: low, lower-middle, upper-middle, and high income, using Gross National Income (GNI) per capita in 2021 (World Bank, 2022).

4 Empirical Strategy

Previous work has mostly compared aggregate net ethnic flows in and out of neighborhoods, usually by decennial censuses (Card, 2001; Card et al., 2008; Saiz and Wachter, 2011). While net outflow-to-inflow parameters can be informative, they have limitations.

Notably, they ignore gross churn and potential sorting among individuals from the same ethnic group. More problematically, reduced *net* outflows of majorities could be mechanically associated with minorities' arrivals, as the housing units occupied by the latter become available (Moraga et al., 2019). Using past immigrant densities to predict future native exits is a better practice (Andersson et al., 2021), but results may still be endogenous to the underlying natives' *propensities to churn.*[17] Note that conventional shift-share instruments cannot address this issue, because first-period immigrant neighborhood choices may depend on the underlying local mobility of the native population and, therefore, on subsequent available housing vacancies.

This problem can be surmounted by modeling the probability of native moves explicitly, allowing for the inclusion of a large set of neighborhood and family characteristics that account for mobility. We thus consider the following linear probability model:

$$P(y_{i,z,t+1} = 1 | \mathbf{X}_{i,z,t}, s_{z,t}, \lambda_{r,t}, \lambda_z) = \alpha + \beta \cdot s_{z,t} + \Theta' \mathbf{X}_{i,z,t} + \lambda_{r,t} + \lambda_z + \epsilon_{i,z,t+1} \quad (11)$$

where $y_{i,z,t+1}$ is a relocation dummy (a dichotomous variable that assumes value one if the individual i moved to another neighborhood in the next year, $t+1$); $s_{z,t}$ represents the share of immigrants from non-Western countries that resided in the same location z (neighborhood or building complex) as the individual i; and $\mathbf{X}_{i,z,t}$ is a set of control variables that include characteristics of the head of family (age, gender, marital status, education level, and employment status), household (disposable income, size, and number of children), housing information (tenure type, the number of rooms in the unit, and building's age), and time-changing amenities (number of banks, hospitals, doctors, primary schools, grocery stores, entertainment businesses, share of employed residents, and restaurants). ϵ_{it} is a random error term. We cluster the standard errors at the neighborhood level in all specifications. While hazard models at this scale are computationally unfeasible, we also control for the household's length of stay in the house, thereby capturing changing conditional probabilities of exit as time passes.

Importantly, we also control for interactions between a five-year trend variable and time-invariant neighborhood characteristics: distance to the city center, to a highway, to the coast or lake, to waste disposal facilities, to forests, and to train stations. These interactions capture long-term trends associated with different neighborhood typologies, more saliently growing suburbanization. Conventional shift-share IV strategies do not typically control for secular trends in the types of neighborhoods where immigrants clustered originally, a problem akin to the potential endogeneity of initial industry shares in

[17] E.g., immigrants may tend to move to places with a disproportionate proportion of young, unmarried individuals with unstable incomes, and short housing tenures. These are also the most likely individuals to move out subsequently.

Bartik instruments (Goldsmith-Pinkham et al., 2020).

In addition, to control for unobservable time- or location-specific factors that affect the decision to move, we also include region-by-year fixed effects,[18] $\lambda_{r,t}$, and neighborhood fixed effects, λ_z. Therefore, the OLS identification is initially based on changes in lagged immigrant shares by neighborhood relative to changes in other areas in the same region and year. In later specifications, where we assess the presence of immigrants in a building complex, we further include neighborhood-by-year and building complex fixed effects.

We see this model as one based on revealed preference: under the null hypothesis of no ethnic preferences, the probability of a native exit should not relate to the presence of non-Western migrants, conditional on counterfactual mobility. Note that by using micro-data, we capture exits that may be countered by corresponding native arrivals in the aggregate.

4.1 Instrumental Variable Design

Even after controlling for individual propensities to move, neighborhood fixed effects, and local varying time trends, there is still a possibility of endogeneity bias due to unobserved shocks. The direction of bias cannot be determined *ex-ante*. For instance, if immigrants are being drawn to neighborhoods with thriving economies–in ways that are not captured by the local employment shares–then the estimated coefficient of interest would be downward biased. Conversely, if immigrants settle in neighborhoods with worsening job prospects, the coefficient could be upward biased.

This leads us to develop a novel approach that builds on the random assignment of newly arriving refugees across municipalities in Denmark. Refugees – see a detailed definition in Appendix B5 – comprised a substantial 27.3% of non-Western immigrants during the period and were less likely to return to their origin countries (Jensen and Pedersen, 2007).

Damm and Dustmann (2014) convincingly argue that the Danish Spatial Dispersal Policy of refugees from 1986-1998 was random across municipalities conditional on a few controls, and Azlor et al. (2020) also documents the random elements of the refugee dispersal policies from 1999 until the present. Concretely, placement agencies situated refugees across municipalities to avoid excessive concentration. Refugees were placed on social or private rentals depending on the random availability of housing units on the week of arrival: Damm (2014) offers evidence that the refugee dispersal policy was random at the neighborhood level, conditional on observed individual characteristics that we also

[18]We adopt the region definition from the 2007 Municipal Reform that created five regional administrative regions above municipalities but below the central government.

include in our specifications.

Nevertheless, refugees have a degree of agency, and their random initial placement does not imply that they stay put forever. Subsequent refugee moves are endogenous and susceptible to generating instruments that violate the exclusion restriction assumption. To address these issues, we develop a novel *Arrival-Stayer Markov Instrumental Variable (ASM-IV)* to purge any endogeneity that could arise from using the *current* share of refugees by neighborhood. Consider the number of *refugees* in neighborhood i in time t, $R_{i,t}$, as the sum of refugees who stayed from the previous year, $S_{i,t}$, plus the refugees moving in from other neighborhoods, $D_{i,t}$, plus the total inflow of new arrivals in the country placed into i, $A_{i,t}$.

$$R_{i,t} = S_{i,t} + D_{i,t} + A_{i,t} \tag{12}$$

We can further separate each of the components above into social housing dwellers, Pb, and refugees living in private housing Pv:

$$\begin{aligned} R_{i,t} &= R_{i,t}^{Pb} + R_{i,t}^{Pv} \\ &= S_{i,t}^{Pb} + S_{i,t}^{Pv} + A_{i,t}^{Pb} + A_{i,t}^{Pv} + D_{i,t}^{Pb} + D_{i,t}^{Pv}. \end{aligned} \tag{13}$$

In what follows, we propose a Markov-chain constructed IV approach that computes predicted local inflows and outflows of refugees today based on new arrivals, the pre-existing distributions of both social housing and refugees in the past, combined with year-specific mechanical transition probabilities.

Arrivals from abroad $(A_{i,t})$ Per the spatial dispersion policy, the allocation of new refugees from abroad into public housing was likely orthogonal to unobserved subsequent neighborhood shocks. Therefore, we assume that $A_{i,t}^{Pb}$ is exogenous.

However, arrivals into private housing may still be susceptible to endogeneity, considering that some landlords may refrain from renting out to refugees. To construct our instrument, we thus replace the actual probability of refugees moving into neighborhood i, with the share of refugees living there in 1987, $\hat{\theta}_{i,1987}^{R}$, multiplied by the total country-level number of new refugee arrivals into private rentals from abroad, but excluding neighborhood i. Therefore, the predicted number of new arrivals from abroad going to private rentals at i becomes:

$$\hat{A}_{i,t}^{Pv} = \hat{\theta}_{i,1987}^{R} \cdot \sum_{j \neq i} A_{j,t}^{Pv} \tag{14}$$

Stayers $(S_{i,t})$ The number of refugees who stay from one year to the next may also be susceptible to endogeneity concerns, considering the potential attitudes of local natives toward them. To address this concern, we compute the predicted number of refugees who stayed in public or private housing in neighborhood i using the following depreciation rule:

$$\begin{aligned} \hat{S}_{i,t}^{Pb} &= \left(1 - \hat{\rho}_{i,t}^{Pb}\right) \cdot R_{i,t-1}^{Pb} \\ \hat{S}_{i,t}^{Pv} &= \left(1 - \hat{\rho}_{i,t}^{Pv}\right) \cdot R_{i,t-1}^{Pv} \end{aligned} \tag{15}$$

where $\hat{S}_{i,t}^{Pb}$ is the predicted number of refugees who stayed in neighborhood i in public housing. $\hat{S}_{i,t}^{Pv}$ represents the predicted stayers in private housing. $\hat{\rho}_{i,t}^{Pb}$ is a moving probability calculated at the national level (but excluding i), corresponding to the empirical share of public housing refugees who moved out of their neighborhoods at t. $\hat{\rho}_{i,t}^{Pv}$ is calculated analogously for refugees living in private housing units at $t-1$.

Arrivals from other neighborhoods within the country ($D_{i,t}$) Resident refugees who changed neighborhoods may be driven by confounding local destination shocks that also impacted the mobility of natives. To address this concern, we use Equation (15) to predict the overall number of refugees moving to a different neighborhood:

$$\begin{aligned} \hat{N}_{i,t,Pb} &= \hat{\rho}_{i,t}^{Pb} \cdot R_{i,t-1}^{Pb} \\ \hat{N}_{i,t,Pv} &= \hat{\rho}_{i,t}^{Pv} \cdot R_{i,t-1}^{Pv}. \end{aligned} \tag{16}$$

$\hat{N}_{i,t,Pb}$ and $\hat{N}_{i,t,Pb}$ are the predicted number of refugees who moved out of i and into a different neighborhood, conditional on hailing from public or private housing, respectively. Now, we calculate $\hat{\delta}_{i,t}^{Pb}$ as the national frequency – excluding i – of refugees moving back into public housing, conditional on having moved out of public housing in the previous period. Analogously, denote $\hat{\delta}_{i,t}^{Pv}$ as the national average frequency – excluding i – for refugees moving to private housing, conditional on having lived in private housing in $t-1$. Now, consider:

$$\begin{aligned} \hat{M}_{i,t}^{Pb} &= \hat{\delta}_{i,t}^{Pb} \cdot \hat{N}_{i,t,Pb} + \left(1 - \hat{\delta}_{i,t}^{Pv}\right) \cdot \hat{N}_{i,t,Pv} \\ \hat{M}_{i,t}^{Pv} &= \hat{\delta}_{i,t}^{Pv} \cdot \hat{N}_{i,t,Pv} + \left(1 - \hat{\delta}_{i,t}^{Pb}\right) \cdot \hat{N}_{i,t,Pb} \end{aligned} \tag{17}$$

where $\hat{M}_{i,t}^{Pb}$ and $\hat{M}_{i,t}^{Pv}$ are the total expected number of refugees who moved out of the neighborhood i, and went on to live in public or private housing elsewhere, respectively. Figure 4 illustrates the Markov probability chain for movers and stayers from or at i.

Endogeneity concerns also arise regarding the reallocation choices of internal refugee movers. To address them, we compute predictions using the pre-existing shares of public

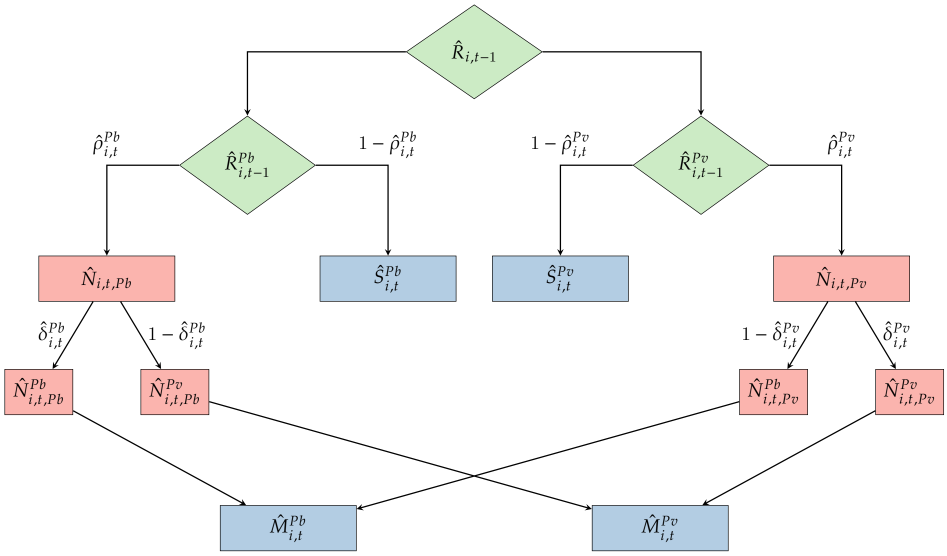


Figure 4. IV Design - Decomposing Refugees into Predicted Stayers and Movers. This figure illustrates the Markov process of refugee immigrants ($\hat{R}_{i,t-1}$) in neighborhood i, into the expected number of refugees that stayed in public housing ($\hat{S}^{Pb}_{i,t}$) or private housing ($\hat{S}^{Pv}_{i,t}$), and those that moved away to public housing ($\hat{M}^{Pb}_{i,t}$), or private housing ($\hat{M}^{Pv}_{i,t}$) in a different neighborhood.

housing or refugees by neighborhood in 1987 as follows:

$$\begin{aligned} \hat{D}^{Pb}_{i,t} &= \hat{\theta}^{Pb}_{i,1987} \cdot \sum_{j \neq i} \hat{M}^{Pb}_{j,t} \\ \hat{D}^{Pv}_{i,t} &= \hat{\theta}^{R}_{i,1987} \cdot \sum_{j \neq i} \hat{M}^{Pv}_{j,t} \end{aligned} \tag{18}$$

where $\hat{\theta}^{Pb}_{i,1987}$ is the national share of public housing units located in neighborhood i in the reference year, 1987. $\hat{\theta}^{R}_{i,1987}$ is the national share of refugees that were living in the neighborhood i in 1987. The intuition is that refugees moving within the public housing system are much more likely to end up in neighborhoods according to the pre-existing distribution of public units. Similarly, refugees moving into private housing are more likely to follow their baseline geographic distribution.

Arrival-Stayer Markov Instrumental Variable (ASM-IV) From Equation (13) and the forecast flows discussed above, we compute the predicted number of refugees in neighborhood i and year t as follows:

$$\begin{aligned} \hat{R}_{i,t} &= \hat{R}^{Pb}_{i,t} + \hat{R}^{Pv}_{i,t} \\ &= \hat{S}^{Pb}_{i,t} + \hat{S}^{Pv}_{i,t} + A^{Pb}_{i,t} + \hat{A}^{Pv}_{i,t} + \hat{D}^{Pb}_{i,t} + \hat{D}^{Pv}_{i,t}. \end{aligned} \tag{19}$$

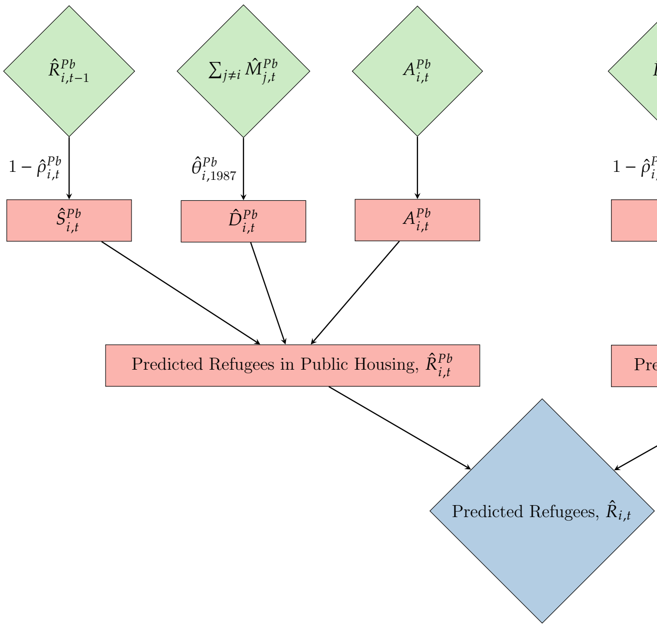


Figure 5. IV Design - Structure Summary. This figure summarizes the structure we adopt to compute the predicted number of refugees in neighborhood i, and year t.

Figure 5 summarizes the structure we use to compute the predicted number of refugees living in each neighborhood at each time. Note that we use 1987 as our reference year, at which we start the predicted series, $\hat{R}_{i,0} = R_{i,1987}$. Denote by $L_{i,t}$ the total number of residents in the neighborhood i and year t. Our final instrumental variable is then given by the predicted share of refugees living in i at time t:

$$\text{ASM-IV}_{i,t} = \frac{\hat{R}_{i,t}}{L_{i,t}} \tag{20}$$

Of course, not every non-Western immigrant entered Denmark as a refugee. However, economic immigrants' location decisions are not randomly assigned. In addition, the national identity of migrants changed substantially during the period: starting from Southern Europe, then the Balkans, Eastern Europe, and Turkey, and more recently from Pakistan, Somalia, Syria, and Iraq. This implies that conventional shift-share instruments based only on the initial location by ethnic group and their national growth tend to be much weaker than their ASM-IV counterparts.

Results using our ASM-IV need to be interpreted in theory as Local Average Treatment Effects (LATE) (Angrist and Imbens, 1995), as responses to local immigration shocks driven by *refugee* inflows. Because refugees are at the center of current political debates this is a relevant LATE. Moreover, in practice, IV results are close to OLS ones, and we explicitly examine treatment effect heterogeneity later on.

4.2 Shocks versus Shares

Our constructed instrument is not directly based on shift-shares, as it combines refugee shocks, national transition probabilities, and the neighborhood shares of social housing and refugee settlement in 1987. Our highly-saturated models are also designed to capture differential local trends. Nonetheless, there remains an "exposure" component to the identification strategy (Goldsmith-Pinkham et al., 2020). Here, we have only two initial shares (akin to sectors in a shift-share analysis): social housing and refugees. Therefore, one cannot exploit the orthogonality of shocks across large numbers of multiple sectors (Borusyak et al., 2021). However, we can still assess how the "shares" component compares to the "shocks" one.

To do so, we decompose our ASM-IV into two orthogonal components: (i) a share-based component and (ii) a residual component based on conditional randomization across neighborhoods (shocks). One may be first tempted to recalculate a Markov process using refugee arrivals as shocks, but randomizing their subsequent moves. However, new refugee arrivals were more likely to be assigned to social housing than the population at large. Therefore, refugee arrival "shocks" still correlate with the 1987 shares of public housing.[19] Therefore, to ensure orthogonality between components, we run the following neighborhood-level specification for each year in our sample separately, situating the ASM-IV on the left-hand side:

$$\text{ASM-IV}_{i,t} = \alpha_t + \beta_{1,t} \cdot \hat{\theta}^{Pb}_{i,1987} + \beta_{2,t} \cdot \hat{\theta}^{R}_{i,1987} + \epsilon_{i,t}. \tag{21}$$

As described in Section 4.1, $\hat{\theta}^{Pb}_{i,1987}$ and $\hat{\theta}^{R}_{i,1987}$ are the public housing share and refugee share in neighborhood i in 1987, respectively. Denoting the estimates of α_t, $\beta_{1,t}$, and $\beta_{2,t}$ by $\hat{\alpha}_t$, $\hat{\beta}_{1,t}$, and $\hat{\beta}_{2,t}$, and the model's residuals by $\hat{\epsilon}_{i,t}$, we compute the following two orthogonal components:

$$\begin{aligned} \text{Shares Component}_{i,t} &= \hat{\alpha}_t + \hat{\beta}_{1,t} \cdot \hat{\theta}^{Pb}_{i,1987} + \hat{\beta}_{2,t} \cdot \hat{\theta}^{R}_{i,1987} \\ \text{Shocks Component}_{i,t} &= \hat{\alpha}_t + \hat{\beta}_{1,t} \cdot \bar{\hat{\theta}}^{Pb}_{1987} + \hat{\beta}_{2,t} \cdot \bar{\hat{\theta}}^{R}_{1987} + \hat{\epsilon}_{i,t} \end{aligned} \tag{22}$$

where $\bar{\hat{\theta}}^{Pb}_{1987}$ is the average public housing share in 1987 and $\bar{\hat{\theta}}^{R}_{1987}$ is the average refugee share across neighborhoods in 1987. We then use both components as instruments and check whether they yield similar results using a Sargan test, as proposed by Goldsmith-Pinkham et al. (2020). Note that the shocks component captures year-specific surprises

[19] This is not surprising as local placement agencies were pressed for time, and local governments could assign up to 1/3 of new vacancies in social housing to high-need families, outside of the waiting list order. While addressing it directly in our paper, we red-flag this issue for future research using refugee shocks in a Scandinavian context.

Table 3. Baseline Regressions

	Dependent Variable: Moved out of the Neighborhood in $t+1$					
	Full Sample		Single Family	Small Building or Complex (less than 10 units)	Large Building or Complex (10 or more units)	
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant share (neighborhood)	0.038***	0.021***	−0.002	0.004	0.036***	
	(0.009)	(0.006)	(0.005)	(0.010)	(0.008)	
Immigrant share (building complex)						0.015***
						(0.004)
N	53,332,175	53,332,175	27,978,943	10,740,642	14,612,584	14,612,240
R^2	0.028	0.092	0.050	0.108	0.084	0.101
Controls		✓	✓	✓	✓	✓
Region × Year FEs	✓	✓	✓	✓	✓	
Neighborhood FEs	✓	✓	✓	✓	✓	
Neighborhood × Year FEs						✓
Building Complex FEs						✓

Notes: The observations consist of heads of households, non-immigrants, with ages above 18 years old. Standard errors in parentheses are clustered by neighborhood. Statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

to immigrant arrivals, thereby eschewing concerns about their persistence (Jaeger et al., 2018).

5 Results

5.1 Baseline Regressions

The results of several specifications based on equation (11) are shown in Table 3, with standard errors clustered at the neighborhood level. In column (1), we control for neighborhood and region-by-year fixed effects using complete data from approximately 53 million observations of household heads over time. The coefficient indicates that an increase in the immigrant share by 30 percentage points, for example, is associated with a 1.14 percentage point increase in the probability of native out-mobility from the neighborhood. However, areas with higher proportions of foreign-born individuals may also have populations with a higher propensity to migrate. Note that neighborhood fixed effects do not fully address this issue because *changes* in immigrant shares—as opposed to their levels—may simply result from increased housing market availability in areas with high native churn.

In column (2), we include household characteristics and time-varying neighborhood attributes. These factors substantially impact the estimates, suggesting that the baseline mobility of natives was indeed higher in high-immigration locations and periods. Hence, future studies of white flight cannot ignore the underlying differences in churn rates that arise from the local characteristics of the majority population.

In quantitative terms, the results still reveal substantial flight. An increase of 30 percentage points in the proportion of non-Western immigrants is associated with a 0.63 percentage point increase in out-migration from the neighborhood in any given year, which is 7.9 percent higher than the baseline. Naturally, the overall likelihood of moving increases over the years.

Column (3) focuses on single-family homes, which are predominantly situated in suburban areas. We estimate a precisely estimated zero effect for native flight from within this group of dwellings.[20] This evidence – consistent with Moraga et al. (2019) and Salazar (2020) – allows us to formulate the hypothesis that the urbanistic characteristics of the neighborhood mediate ethnic displacement. By distancing people and hindering unwanted social interactions, suburban environments might reduce social tensions across groups. Future research on ethnic displacement may want to distinguish between suburban and denser urban areas and test this hypothesis in other environments. In column (4), we focus on attached single-family homes or small buildings with fewer than 10 apartment units. Similarly to column (3), we find a precise zero estimate for the coefficient on the share of non-Western immigrants within this group.

Conversely, column (5) shows strong results within the set of large multifamily complexes, hosting about one-third of households in the country.[21] In that subsample, a 30 percentage point increase in the share of immigrants in the neighborhood increases the probability of native exits by 1.08 percentage points. This suggests that building and neighborhood typologies that allow for greater daily interactions between immigrants and natives are also associated with stronger native flight. This result contrasts with studies arguing that personal exposure may generate better natives' attitudes towards the foreign-born (Bursztyn et al., 2024; Andries et al., 2023). However, the evidence may be reconciled by the fact that most interactions between people living in very dense areas are anonymous.

Given these findings, the specification in Table 3, column (5) constitutes our baseline for the study of heterogeneous treatment effects below, as we focus on better understanding native flight *where it is a relevant phenomenon.*

5.2 Dealing with Unobservable Shocks: ASM-IV

Panel A in Table 4 presents the results using the arrival-stayer Markov predictions as instruments. Controls for neighborhood fixed effects, region-by-year fixed effects,

[20] Although there could be flight from (into) there and into (out of) other areas.

[21] Note again that, quite demandingly, we are eliminating the variation *between* large and small units.

neighborhood-type trends, changing amenities, and evolving household characteristics are included throughout.

Our 2SLS estimates have similar magnitudes to the baseline OLS estimates. In all cases, the first-stage F-statistics are large. These specifications suggest that the results are not driven by unobserved shocks, simultaneously attracting immigrants and repelling natives.[22] Considering the similarity between the 2SLS and OLS coefficients, we adopt the latter specification (column (5) of Table 3) in further analyses.

In Panel B of Table 4, we include the orthogonal "shocks" and "shares" components of ASM-IV as separate instruments. Overall results are unchanged. Using the full sample, columns 1 (uncontrolled) and 2 (all controls), a Sargan test cannot reject the equality between coefficients using the two alternative orthogonal sources of variation. These results strongly validate the ASM-IV approach.

Conditioning on multifamily buildings (column 3) yields stronger native-flight results, but we reject the hypothesis of strict equality between the two instruments. This occurs because the "shocks" instrument is rather weak on its own. It is arguably too much to expect its variation to be substantial in such an oversaturated model, which uses only one-fourth of the sample and variance between large buildings. Although we are confident that the instruments capture exogenous variation across neighborhoods, in the following section, we use an alternative source of variation to demonstrate the robustness of the results within the subsample of large buildings.

5.3 Neighborhood-by-year Fixed Effects

As in previous studies, we have thus far used the variance in ethnic composition across neighborhoods. However, having an immigrant neighbor immediately next door may be as important for native behavior. We now return to Table 3, focusing on column 6, to leverage the richness of the data and investigate additional impacts at very local interaction levels. Specifically, we focus on the 616,094 homes in buildings with more than ten units.[23] To estimate the effects that go beyond those captured by immigrant neighborhood shares, we now use both neighborhood-by-year and building-complex fixed effects.

[22]In Table D.3 in the Appendix, we also tested the sensitivity of our OLS estimates to the inclusion of two additional interaction terms between a five-year trend variable and: (i) the 1987 neighborhood shares of refugees and (ii) the 1987 neighborhood share of public housing. The idea is that these interaction terms capture the long-term trends in moving rates associated with the 1987 distribution of public housing and refugees across neighborhoods in Denmark. The estimates in Table D.3 are very similar to the ones in Table 3, suggesting that our results are robust to controlling for the trends associated with the exposure measures we use to build the ASM-IV.

[23]Note that estimates of the immigrant share become very volatile for buildings with nine or fewer housing units. However, results are still robust to different cutoffs for the minimum size of a complex (Table D.2 of Appendix D2).

Table 4. Two-Stage Least Squares Results

	Dependent Variable: Moved out of the Neighborhood in $t+1$		
	Full Sample		Large Building or Complex (10 or more units)
	(1)	(2)	(3)
Panel A. ASM-IV			
Immigrant share (neighborhood)	0.032**	0.024**	0.046***
	(0.014)	(0.010)	(0.015)
N	52,031,888	52,031,888	14,269,725
First Stage Results			
Arriver-Stayer-Markov Instrumental Variable (ASM-IV)	1.883***	1.859***	1.832***
	(0.087)	(0.087)	(0.113)
First Stage F-statistic	471.163	460.938	261.449
N	52,031,888	52,031,888	14,269,725
Panel B. Orthogonal Decomposition			
Immigrant share (neighborhood)	0.048***	0.032***	0.055***
	(0.015)	(0.010)	(0.016)
N	52,031,888	52,031,888	14,269,725
First Stage Results			
Shares Component	1.793***	1.762***	1.612***
	(0.096)	(0.096)	(0.132)
Shocks Component	0.687***	0.699***	1.249***
	(0.143)	(0.145)	(0.198)
First Stage F-statistic	204.868	198.936	124.167
Sargan Test/Hansen J-statitic (p-value)	0.157 (0.692)	0.001 (0.971)	4.625 (0.032)
N	52,031,888	52,031,888	14,269,725
Controls		✓	✓
Region × Year FEs	✓	✓	✓
Neighborhood FEs	✓	✓	✓

Notes: The observations consist of heads of households, non-immigrants, with ages above 18 years old. Standard errors in parentheses are clustered by neighborhood. Statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Panel A presents the results where we instrument the share of immigrants in a neighborhood using the ASM-IV, as described in Section 4.1. Panel B shows the results when we use the orthogonal shares and shocks components as described in Section 4.2.

One advantage of saturating the model so aggressively is that the neighborhood-by-year fixed effects should capture any other time-varying, unobservable local shocks impacting native mobility. Note that this specification uses a very different source of variation than in earlier estimates, which were based on differences *between* neighborhoods. We are now studying native exits by comparing different buildings *within* the same neighborhood and time period. Although the definition of the treatment is different, identifying native flight here should lend substantial credibility to the existence of this phenomenon.

Even after soaking up a large amount of the variation in the data, we still observe *additional* effects of the micro concentration of non-Westerners on native exits across buildings in the same area and year. Consider, for instance, two housing complexes within the same neighborhood. One has a 30 percent foreign-born population, while the other is entirely indigenous. On average, there will be a 0.45 percentage point higher probability of native exits from the former. Note that this impact is in addition to any neighborhood-level effects. Thus, ethnic interactions appear to layer up at different geographic extents.

5.4 Robustness Tests

We conduct several robustness tests detailed in Appendix D. First, to ensure that the results are not sensitive to the neighborhood definition, we re-estimate our models using smaller geographic boundaries (Appendix D1). Second, we test alternative cutoffs for defining building complexes, to address concerns about sensitivity to different definitions (Appendix D2). Third, we include additional trend controls that are interacted with the 1987 baseline shares of refugees and public housing. This accounts for potential long-term demographic trajectories that could confound our identification strategy (Appendix D3). Fourth, we exclude neighborhoods that were targeted by Denmark's "Ghetto Programme" to ensure our estimates are not driven by policy responses (Appendix D4). Finally, we address concerns about unobserved building characteristics by recalculating immigrant exposure measures that exclude each household's building. (Appendix D5). Our core findings remain unchanged across all specifications, with native flight effects consistently estimated.

5.5 Non-linearities

Our baseline specifications assume a linear relationship between the growth of the immigrant population and the probability of natives leaving. However, the propensity of locally born citizens to move away could accelerate in areas with substantially higher concentrations of foreign-born individuals. (Schelling, 1971; Card et al., 2008). In this section, we evaluate the nonlinearity of native flight. Our analysis is restricted to building complexes with more than ten units, where the flight phenomenon is stronger.

We start by estimating an individual mobility regression on all variables in equation (11), including the set of fixed effects, but excluding immigrant shares. We then compute the adjusted probability of moving out by adding the residuals back to the predicted probability of exit *at national average characteristics.* Next, we calculate the local averages of these adjusted probabilities — by neighborhood or building complex — and group them into bins according to their lagged shares of immigrants. Finally, we estimate a continuous, smooth, restricted cubic spline that models the local means of the compositionally adjusted probability of moving out as a function of the average immigrant density in the previous year. We weight the estimates by the number of locations — neighborhoods or building complexes — within each bin.

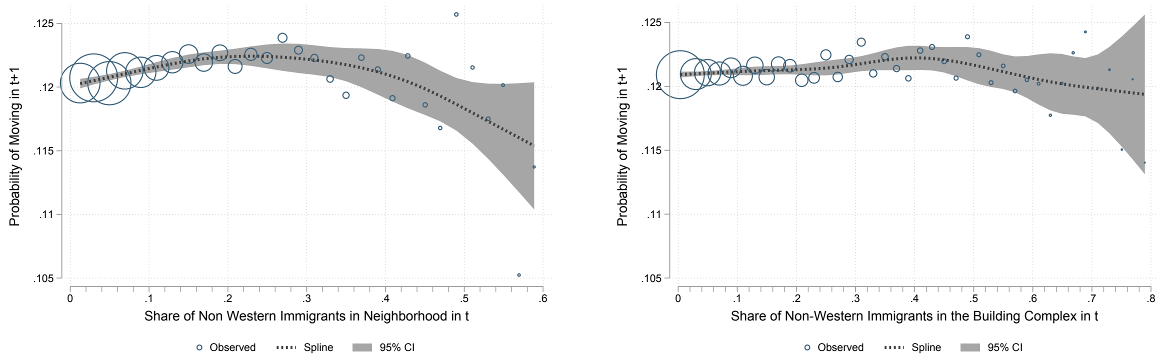
To the best of our knowledge, this specification is novel to the literature. Previous work has primarily focused on net outflow-to-inflow relationships, potentially missing turning points in "white" exits that are compensated by countervailing "white" arrivals. In other words, conventional tipping point estimates completely overlook the spatial sorting of majorities based on their dislike of minorities. Using our approach to find a tipping point in out-mobility, as opposed to net changes in the white share, provides us with weaker sufficient conditions to prove nonlinearities in native ethnic tastes.[24]

Figure 6 shows the estimated smoothed lines with their 95% confidence intervals in gray. The blue circles are centered on local native mobility averages, and their sizes are proportional to the number of observations within each bin. Panel A illustrates the relationship at the neighborhood level, and Panel B shows the relationship for building complexes. Note that the latter figure is purged of neighborhood-by-year fixed effects.

Panel A shows a mildly inverted U-shaped relationship between the percentage of immigrants and the likelihood that a native family will move out of its neighborhood. This likelihood increases with immigrant density up to around 30%, but then decreases to the baseline of 43%. This is consistent with the idea that there is a distribution of native tolerance (or moving costs). Less accommodating households (or those with low mobility costs) leave first, until the average native is indifferent to the neighborhood's composition (or faces high mobility costs).

In other words, at a non-Westerner share of about 43%, those who wanted to move out disproportionately already have. Further neighborhood ethnic change may still happen, but it must now occur through replacements in naturally occurring vacancies. However, confidence intervals also increase substantially at high foreign-born shares, so any

[24]However, not necessary conditions, as majority population *inflows* might still depend nonlinearly on the immigrant share, even if *outflows* did not.



(a) Neighborhood Exposure

(b) Building Complex Exposure

Figure 6. Non-linear relationship between the share of immigrants and the move-out probability. This figure presents a local average smooth plot of adjusted native move-out probability by the lagged share of immigrants in the neighborhood (Panel A) or the building complex (Panel B). The shaded areas represent the 95% confidence intervals for coefficients.

conclusions here must be tentative. In any case, there is no clear tipping point.

Panel B shows a similar pattern for the percentage of immigrants in the building complex, with a turning point at around 40%. However, this highly saturated model becomes imprecise at high exposure rates. Therefore, we conclude that there is no evidence of tipping point dynamics whereby native exits accelerate after crossing a threshold.[25]

5.6 Heterogeneity of Native's Responses

Next, we study treatment effect heterogeneity: Which types of natives are more likely to leave? We group natives ex-ante according to major, salient, observable characteristics and test whether a particular group is more sensitive to the presence of immigrants than others are. Specifically, we estimate the effect of the neighborhood share of immigrants, as in Table 3, column 5, but now interacted with dummies for each group:

$$\begin{aligned} P(y_{i,z,t+1} = 1 | \mathbf{X}_{i,z,t}, s_{z,t}, \lambda_{r,t}, \lambda_z) =& \alpha + \Sigma_{g=1}^{k} \beta_g \cdot s_{z,t} \cdot \mathbb{1}(G_i = g) \\ &+ \Sigma_{g=1}^{k} \mathbb{1}(G_i = g) + \Theta' \mathbf{X}_{i,z,t} + \lambda_{r,t} + \lambda_z + \epsilon_{i,z,t+1} \end{aligned} \tag{23}$$

where $\mathbb{1}(G_i = g)$ is a binary variable that equals one if the individual i belongs to group g, based on the socioeconomic characteristics detailed below. The remaining variables are as in Table 3. The specifications include region-by-year and neighborhood fixed effects

[25] Davis et al. (2025) present evidence that challenges prior findings on racial tipping. Using U.S. Census data from 1970 to 2000, they demonstrate that spatial racial spillovers influence racial clustering and the spatial location of racial change. Like us, they argue that integration policies must consider neighborhood heterogeneity.

and are restricted to building complexes with more than ten units. Standard errors are clustered at the neighborhood level. To avoid type I errors or publication biases, Section 5.6.4 shows that significant results can be reproduced using alternate random split samples.

5.6.1 Heterogeneity by Age, Income, and Education

First, we group natives according to their age, education level, and income. Household heads under the median age of 47 are classified as young; otherwise, they are classified as old. We also classify households as low- or high-income depending on whether their disposable income was below or above the median for that year. Individuals are classified as having a low level of education if their highest level of education was primary school (38 percent of the sample), and as having a high level of education if they completed vocational training or further education. Figure 7 plots the estimated coefficients for the interaction terms. We find that older heads of households tend to move more in response to non-Western immigrants, and this tendency is consistent across different levels of education and income. There are no statistically significant effects for young heads of households, regardless of their income or education level.

Our findings show that age is an important driver of flight behavior among natives. One potential explanation for why age plays such a critical role is its association with political preferences. Recent European surveys show that younger adults tend to lean left politically and hold more positive views toward immigration than older age groups do (Silver and Johnson, 2018). These differences in political preferences may result in different behavioral responses when confronted with neighborhood demographic change. Older natives demonstrate lower tolerance for the presence of immigrants and consequently a higher propensity for residential relocation.

5.6.2 Treatment Effects by Age, Income, and Parental Status

We now consider that households with children may have unique constraints or preferences regarding local amenities. In the US, Cascio and Lewis (2012) find that parents' school choices may drive the ethnic segregation of foreign-born children. Following the previous section's approach, we estimate a specification in which we replace educational levels with a binary variable that assumes a value of 1 for households with children under 18 years old.

Figure figure 8 plots the coefficients, which show again that older heads of households are more likely to exit. In contrast, the presence of children in the household does not appear to increase the likelihood of leaving. One potential explanation for this result is

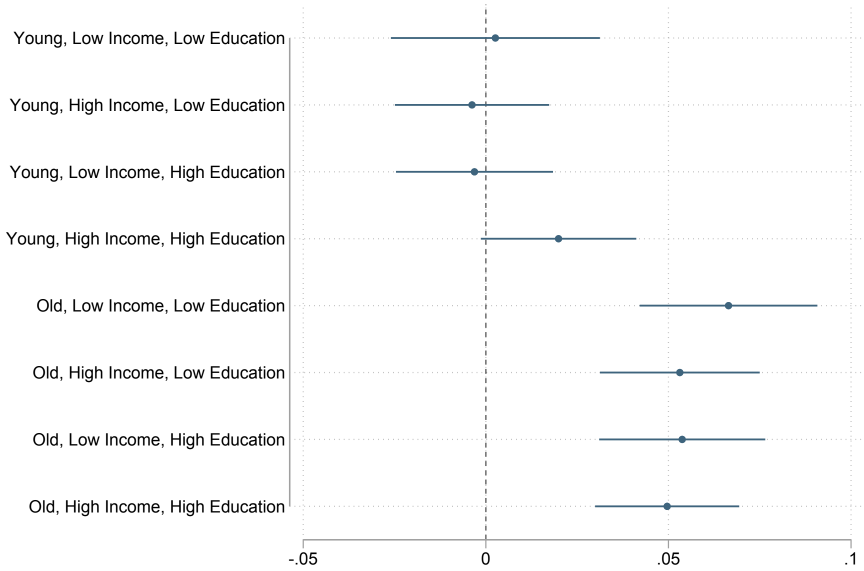


Figure 7. Heterogeneity by Age, Education, and Income. In this figure, we plot the estimated coefficients for the heterogeneity analysis as described in Section 5.6.1. The coefficients are the interaction terms of the share of non-Western immigrants in the neighborhood with binary variables corresponding to eight groups. Young are defined as people being younger than 47 years old, the median age in the sample. Low education is defined as the basic level. Natives are also divided according to their household income, above or below the median. The bars represent the 95% confidence intervals for coefficients. Standard errors are clustered at the neighborhood level.

that children may be relocated to other public or private schools rather than the whole family moving to a different neighborhood (Farre et al., 2018; Bjerre-Nielsen and Gandil, 2024).

5.6.3 Treatment Effect Heterogeneity by Housing Tenure

Existing research indicates that homeowners tend to move less often (Jia et al., 2023). Therefore, migration costs may mediate reactions to local ethnic shocks. Another important characteristic is whether families live in social housing, which represents a significant portion of the housing stock in Denmark: 42% of *rental* units (Figure B.4). Research in both Denmark (Munch and Svarer, 2002) and the Netherlands (De Graaff et al., 2009) has established a lock-in effect: households face substantial costs when moving out of social housing because rents are below market rates and the quality of the units is generally good.

Current social housing renters in Denmark are given priority on waiting lists for other units administered by the same housing organization, especially in cases of divorce or childbearing. However, regardless, waiting lists in major cities are years long: in April 2024, 1,095 units — mostly in remote areas — were available, while 601,862 people were

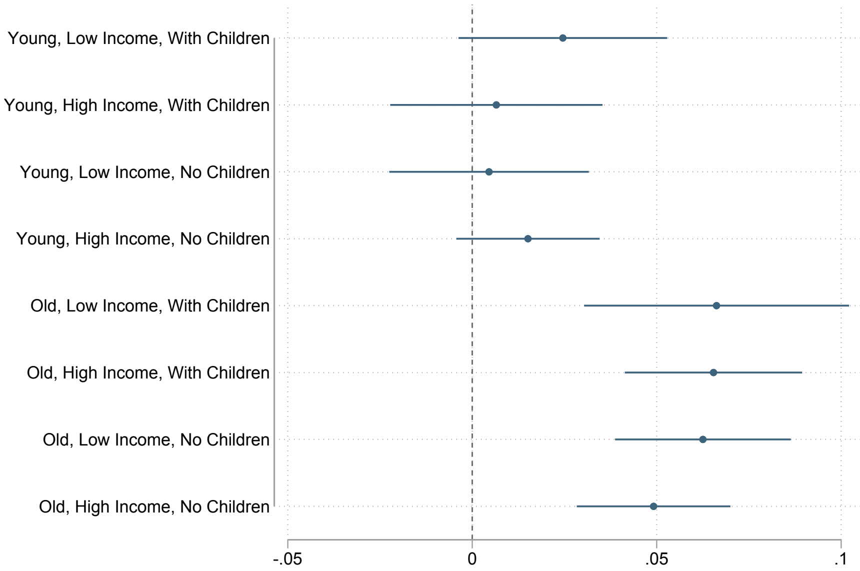


Figure 8. Heterogeneity by Age, Income, and Children in the Household. In this figure, we plot the estimated coefficients for the heterogeneity analysis as described in Section 5.6.2. The coefficients are the interaction terms of the share of non-Western immigrants in the neighborhood with binary variables corresponding to eight groups. Young are defined as people being younger than 47 years old, the median age in the sample. Natives are divided according to their income, above or below the median. Households are also grouped among those with and without children. The bars represent the 95% confidence intervals for coefficients. Standard errors are clustered at the neighborhood level.

on waiting lists across the country.[26] Therefore, it is difficult to swap a social unit for another one.

We thus estimate Equation (23) defining groups as (i) renters living in private housing, (ii) renters living in social housing, or (iii) homeowners. Figure 9 plots the estimated coefficients for the interaction terms, β_g.

We find a negative coefficient for renters of private apartments. However, this estimate does not survive our split-sample robustness test and likely arose by chance. The coefficient for condo homeowners is not statistically significant. A robust finding, however, is the strong native flight effects for renters in social housing. This is surprising given the high opportunity costs of exit for these households.

On the other hand, much of the political discussion in the country has focused on a few prominent public housing "ghettos." While public perceptions are not always empirically accurate and may be exaggerated, our evidence is consistent with the existence of ethnic tensions in the social housing sector. However, according to these results, issues regarding ethnic segregation in social housing are as much a product of native aversion as they are of the tendency of low-income, foreign-born families to cluster together. This suggests that policy approaches cannot focus solely on the latter but must also address the former.

[26]Data accessed from https://www.danmarkbolig.dk/en/ on April 4th, 2024.

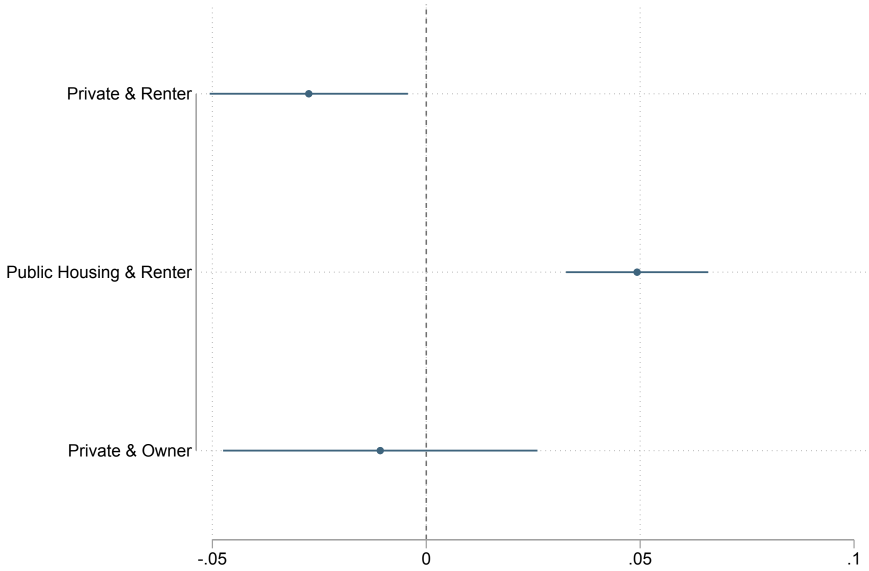


Figure 9. Heterogeneity by Housing Tenure. This figure compares the effects of the immigrant presence on the propensity to move for three groups: public housing renters, private unit renters, and private unit homeowners. The dots are the estimates for the interaction terms of the share of non-Western immigrants in the neighborhood with binary variables corresponding to the defined groups. The bars represent the 95% confidence intervals for coefficients. Standard errors are clustered at the neighborhood.

5.6.4 Multiple Testing

Because we have performed several hypothesis tests, it is natural to be concerned that some of the observed significant results represent false rejections of the null hypothesis. In these situations, conventional Bonferroni-style adjustments can be performed.

However, since we have an abundance of data, it is more convincing to randomly split the sample into two sets of neighborhoods and run the same set of regressions on each set separately (Anderson and Magruder, 2017).[27] Conventional adjustments (Benjamini et al., 2006; Romano and Wolf, 2005) are based on joint hypotheses encompassing all parameters.[28] Unappealingly, such adjustments depend on the *reported* number of coefficients estimated by the researcher. Split-sample approaches validate the parameter of interest more efficiently and are independent of the purported number of parameters being tested. By replicating the results in two random and independent samples, the probability of a Type I error under the null hypothesis is reduced to 0.0025.

[27] We select neighborhoods into each group using alternating positions based on sorting by their identification number (one in, one out, and repeat).

[28] For example, they might set a 5 percent confidence level for the probability that any of the estimators in a group will suffer from a Type I error. Since we generally do not know which of the individual parameters was incorrectly deemed significant in the separate regressions, such adjustments must be very conservative. This is especially true the more parameters we estimate. Therefore, Type II errors for each parameter can be larger than desirable, especially as the number of tests increases.

Figures D.1, D.2 and D.3 in Appendix D6 Show the estimated coefficients for the share of immigrants and their interaction terms with each of the aforementioned groups, for the two random partitions of the sample. The results confirm that emigration by natives is mainly driven by older household heads, while income, education, and the presence of children appear to be irrelevant factors. The response of private renters is the only non-robust coefficient because it does not replicate in both samples.

6 Heterogeneity by Immigrant Type

The previous findings raise a fundamental question about the motivations underlying native households. Thus far, the analysis has considered immigrants from non-Western countries as a uniform group. However, non-Western immigrants have diverse cultural, economic, social, and ethnic backgrounds. One might ask whether native migration decisions are predominantly guided by economic considerations or influenced by biases and ethnic-cultural disparities. For instance, Molla et al. (2022) provides empirical evidence of discrimination in the Swedish private rental housing market based on the names of apartment seekers, particularly those with Muslim-sounding names.

In this section, we evaluate how differences in the characteristics of immigrants may affect the estimates. First, we categorize immigrants by their household income. Those above the median are classified as high-income and those below are classified as low-income. Then, we split the immigrants into four groups based on four different cultural and socioeconomic characteristics. To test the role of religion, we create a binary variable equal to one if the immigrant is from a country where the majority of the population is Muslim, and zero otherwise. Immigrants are allocated into a high language proximity group if they hail from countries with a similarity index to Denmark above the median, or into a low language proximity group otherwise. We assign high social connectedness to immigrants from countries with Facebook connectedness indexes above the median relative to Denmark and low social connectedness otherwise. Finally, we use the World Bank's country income classification groups to separate countries of origin with low or middle-low income from those with middle-high or high income. We then estimate the following specifications:

$$P(y_{i,z,t+1} = 1|\mathbf{X}_{i,t}, s_{z,t}, \lambda_{r,t}, \lambda_z) = \alpha + \Sigma_{g=1}^{4}\beta_g \cdot s_{g,z,t} + \Theta'\mathbf{X}_{i,t} + \lambda_{r,t} + \lambda_z + \epsilon_{i,z,t+1} \quad (24)$$

where $s_{g,z,t}$ represents the share of non-Western immigrants belonging to group g in the number of residents in neighborhood z and year t. The other variables are defined as in Section 4. We group immigrants into four categories using household income and one

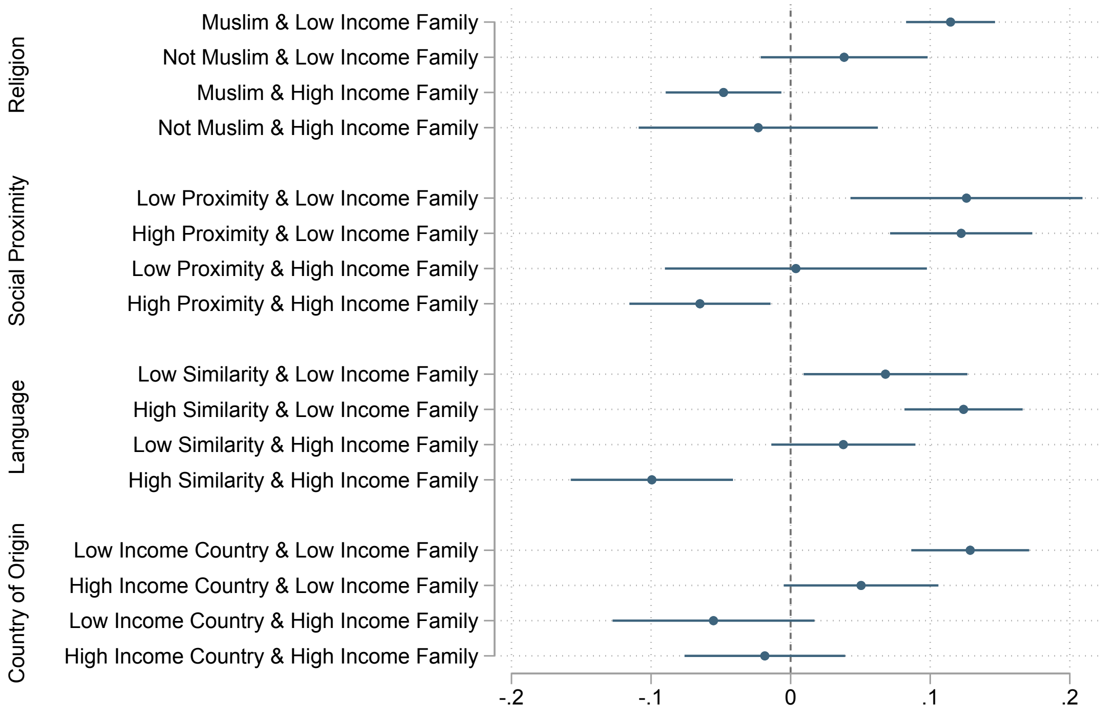


Figure 10. Heterogeneity by Immigrant Group. In this figure, we plot the estimated coefficients for the heterogeneity analysis as described in Section 6. We estimate four regressions where we separate non-Western immigrants based on their household income level and into categories based on their country of origin's religion, social proximity, language similarity, or country income. The bars represent the 95% confidence intervals for coefficients. Standard errors are clustered at the neighborhood level.

of the above-described measures in each of four separate regressions. To focus on areas where flight is relevant, we restrict the sample to building complexes with more than ten units. Finally, we cluster standard errors by neighborhood.

The first four rows of Figure 10 show parameter estimates and confidence bands for Muslim immigrants. The presence of low-income immigrants from Muslim countries positively and significantly affects native exits, while high-income Muslim immigrants reduce them. On average, a one percentage point increase in the proportion of low-income immigrants from Muslim countries in a neighborhood increases the probability of a native moving out by 0.13 percentage points.

Regardless of language proximity, low-income immigrants are associated with flight. Estimates for groups based on Facebook similarity or the wealth of the country of origin show the same pattern. Evidence shows that, rather than being based on religion, language, culture, or national origin, native flight in Denmark seems to be driven by the avoidance of foreign-born individuals with low socioeconomic status. These results are consistent across random split samples, as can be ascertained in Figure D.4 in Appendix D6).

7 Incoming Residents and Native Sorting

Mobility responses to ethnic shocks are sufficient to demonstrate the presence of in-group preferences. However, long-term neighborhood outcomes also depend on who moves in to fill vacancies as they arise disproportionately.

7.1 Immigrant Snowballing

First, we investigate potential foreign-born snowball patterns of settlement. We then restrict our analysis to a subsample that only includes vacant housing units. Since our data is collected annually, a unit is considered vacant if the reference householder is living in a different unit at time $t+1$. Then, we categorize the incoming head of the vacant unit at time $t+1$, or as soon as the unit becomes occupied again. Our goal is to learn who is moving into areas with high minority concentrations. To achieve this, we estimate the following linear probability model:

$$P(M_{i,z,t+1} = 1 | y_{i,z,t+1} = 1, \mathbf{X}_{i,z,t}, s_{z,t}, \lambda_{r,t}, \lambda_z) = \alpha + \beta \cdot s_{z,t} + \Theta' \mathbf{X}_{i,z,t} + \lambda_{r,t} + \lambda_z + \epsilon_{i,z,t+1} \tag{25}$$

where $M_{i,z,t+1} = 1$ is a dichotomous variable assuming value 1 if a non-Western immigrant moved into a vacant address i. The rest of the variables and fixed effects are as in Table 3, and we cluster the standard errors at the neighborhood level. Table 5, column (2) shows that a 10 percentage point increase in the minority population of a neighborhood increases the probability that a non-Western immigrant will move into a vacant unit by 5 percentage points. Similar results are obtained for multifamily buildings (column 5). These effects are very large, as the national average of non-Westerners in large buildings in the nation is of only 6 percent. However, most households that move in are still headed by a native, even in neighborhoods with large minority populations.[29]

Column (6) controls for neighborhood-by-year fixed effects and focuses on the immigrant share by building. Note that these effects are additive to those at the neighborhood level. We still observe disproportionate immigrant clustering at such micro-geographies. A building with an additional 10 percent non-Westerners is about 2 percentage points more likely to receive an immigrant after a vacancy opens up.

Note that these results alone do not imply native avoidance. They could simply reflect immigrants' preference for clustering together. In this case, non-Westerners would

[29] For example, in neighborhoods where 80 percent of residents are immigrants, the probability of an immigrant moving in after a vacancy opens is 40 percentage points higher than the national average of 6 percent. This implies that approximately 54 percent of the households that move in are headed by native Danes.

be outbidding natives for their most desired locations near other migrants.

Table 5. Probability of an Immigrant Moving in

	Dependent Variable: Non-Western Immigrant moved in $t+1$					
	Full Sample		Single Family	Small Building or Complex (less than 10 units)	Large Building or Complex (10 or more units)	
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant share (neighborhood)	0.513*** (0.022)	0.502*** (0.022)	0.471*** (0.052)	0.291*** (0.023)	0.514*** (0.024)	
Immigrant share (building complex)						0.187*** (0.010)
N	6,133,589	6,133,589	1,955,179	1,866,274	2,312,063	2,309,763
R^2	0.066	0.070	0.031	0.026	0.081	0.136
Controls		✓	✓	✓	✓	✓
Region × Year FEs	✓	✓	✓	✓	✓	
Neighborhood FEs	✓	✓	✓	✓	✓	
Neighborhood × Year FEs						✓
Building Complex FEs						✓

Notes: The observations consist of heads of households with ages above 18 years old. Standard errors in parentheses are clustered by neighborhood. Statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

7.2 Sorting by Native Avoidance

We showed earlier that the selection of native leavers was not random. An additional– and hitherto unexplored – topic is whether subsequent native-born in-movers also self-select by type. Even in the absence of tipping, sorting processes–whereby locally-born arrivals significantly differ from departing ones–can have significant social implications.

To investigate this issue, we will now focus on the subsample of non-immigrant families who moved into a new home during the data period. Conditioning on native moves is necessary, because we already know from Table 5 that their unconditional probabilities to move into immigrant-dense neighborhoods are smaller. We then run similar specifications as in Equation (25), where the dependent variables are, respectively, four indicators for whether any new native resident householder in a vacant unit is young (column 1), low-income (column 2), low-educated (column 3), or has children (column 4). The definitions for each group follow from Section 5.6.

Table 6, shows that higher local shares of non-Westerners are associated with an increased probability of the arrival of Danish-born families who are disproportionately young, have low incomes, and do not have children. We do not find statistically significant coefficients for educational levels.

The results regarding the age of migrant avoiders are similar to those pertaining to flight. However, the presence of children is now strongly associated with native avoidance.

Table 6. Sorting of New Native Residents

	Dependent Variable: New Native Resident is...			
	Young	Low-Income	Low-Educated	With Children
	(1)	(2)	(3)	(4)
Immigrant share (neighborhood)	0.052***	0.133***	–0.016	–0.291***
	(0.020)	(0.013)	(0.008)	(0.032)
N	1,488,459	1,488,459	1,488,459	1,488,459
R^2	0.092	0.544	0.232	0.1936
Controls	✓	✓	✓	✓
Region × Year FEs	✓	✓	✓	✓
Neighborhood FEs	✓	✓	✓	✓

Notes: The observations consist of Danish heads of households above 18 years old living in large building complexes with ten or more units. Standard errors in parentheses are clustered by neighborhood. Statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Consistent with the model, this can be explained by the high cost of removing children from their social environment, even in places where the share of immigrants is growing. In contrast, once a family has decided to move–and moving costs are sunk–they prefer to avoid neighborhoods with a high concentration of migrants. This asymmetry between flight and avoidance among families with children could be further exploited in future research designs.

These results may also be partially driven by the structure of social housing benefits in Denmark (20 percent of the housing stock). Once a family with children has moved into a social housing unit, it is not feasible for them to give up the unit and return to the waiting list. However, it is straightforward to simply select a more desirable housing complex when enrolling on the waiting list in the first place. The relationship between social housing and immigrant arrivals has been studied in France by Verdugo (2016); Verdugo and Toma (2018), but future research in this area is needed, including the study of native behavior and relationships between natives and migrants within public housing complexes.

The disproportionate arrival of young low-income native households deserves attention as well, since these families *might* provide less positive spillovers and be worse equipped to handle cultural differences. We flag this issue for further research, noting that conventional outflow-to-inflow studies have typically ignored the taxonomies of "white" outmovers, stayers, and new arrivals throughout the ethnic segregation process.

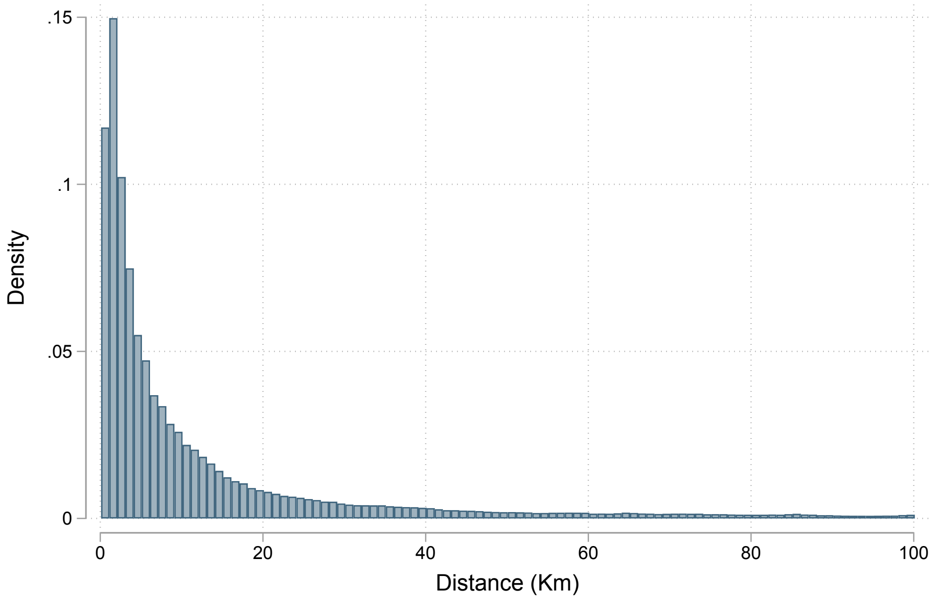


Figure 11. Histogram of Distances Between Origin and Destination Neighborhoods. In this figure, we plot the histogram of distances in kilometers between the origin and the destination neighborhoods for all movers in our data with distances below 100 kilometers. Figure B.5 in the Appendix shows the complete distribution of moving distances.

8 Flight to Where?

We now focus on the ethnic characteristics of the destinations for families moving away from minority areas. The major challenge in this regard stems from mean reversion. According to Table 1, the average native head of household in our data lives in a neighborhood where 4.4% of the residents are non-Western immigrants. Therefore, even if everyone leaving a migrant-dense area were to move at random, we would still measure lower foreign-born shares in destination neighborhoods.

However, people do not move randomly. In fact, moving distances are well characterized by an empirical gravity equation. Figure 11 plots the histogram of distances between the origin and destination neighborhoods for movers in our data. We condition on moves below 100 km for visibility. When people change residences, they tend to move disproportionately to nearby neighborhoods.

Therefore, we use this strong empirical pattern to generate suitable counterfactuals for mobility. Concretely, we define the *Reverting Weighted Average Immigrant Share* (RWAIS) for neighborhood z, as a weighted average of the immigrant shares in all other neighborhoods j, where all weights add up to one and are based on a gravity equation:

$$\text{RWAIS}_{z,t} = \sum_{j \neq z}^{N} \frac{s_{j,t}}{d_{jz}^{\beta} \left(\sum_{j \neq z}^{N} \frac{1}{d_{jz}^{\beta}} \right)} \tag{26}$$

Here $s_{j,t}$ is the share of immigrants in neighborhood j and time t. d_{jz} is the distance between neighborhoods j and z. β represents a spatial decay parameter, which regulates how fast the weights converge to zero. We don't have strong priors for this parameter *in the absence of immigration*, so we first adopt the Newtonian assumption ($\beta = 2$).[30] The purpose of the sum term in the denominator, in the parenthesis, is to re-scale the weights so they add to 1.

Figure 12 is based on the sample of Danish natives who were living in large building complexes and moved to a different neighborhood in the following year.[31] We plot both the potential RWAIS (red line) and the average non-Westerner share at the chosen destination (blue line) – both in the y-axis – against 100 equal-sized bins capturing the mean share of immigrants in the origin neighborhood. In gray, we show the 45-degree line.

There are three main takeaways. First, the blue line shows a positive correlation between immigrant shares at the origin and destination. Second, as the share of immigrants increases, the slope of the curve becomes flatter. However, these basic patterns can be easily explained by the distribution of immigrant shares in proximate neighborhoods, as shown by the counterfactual RWAIS red line. People living in central Copenhagen who need to move to similar areas within the city will naturally be exposed to higher immigration shares than in the rest of the country. As pointed out by Blair (2023), outside options are important to fully understand ethnic segregation.

Most importantly, our third takeaway is that actual exposure for locals in destination neighborhoods is below the counterfactual. Furthermore, the gap between the observed and predicted immigrant shares in the destination area grows as the immigrant share in the origin area increases. Therefore, we conclude that natives moving out of immigrant-dense neighborhoods are looking for new homes in distinctly native areas.[32]

[30] Using the data from Figure 12, the coefficient of a regression of the log of the probability of a move on the log of the distance between origin and destination – omitting the constant – is of -1.83, very close to our baseline choice.

[31] This is where we mostly find white flight. The figures are virtually identical when including all Danish residents, and not just those living in large building complexes. The figures are available upon request.

[32] Appendix D7 shows that the main takeaways discussed here are robust to assuming different values for the spatial decay parameter, β.

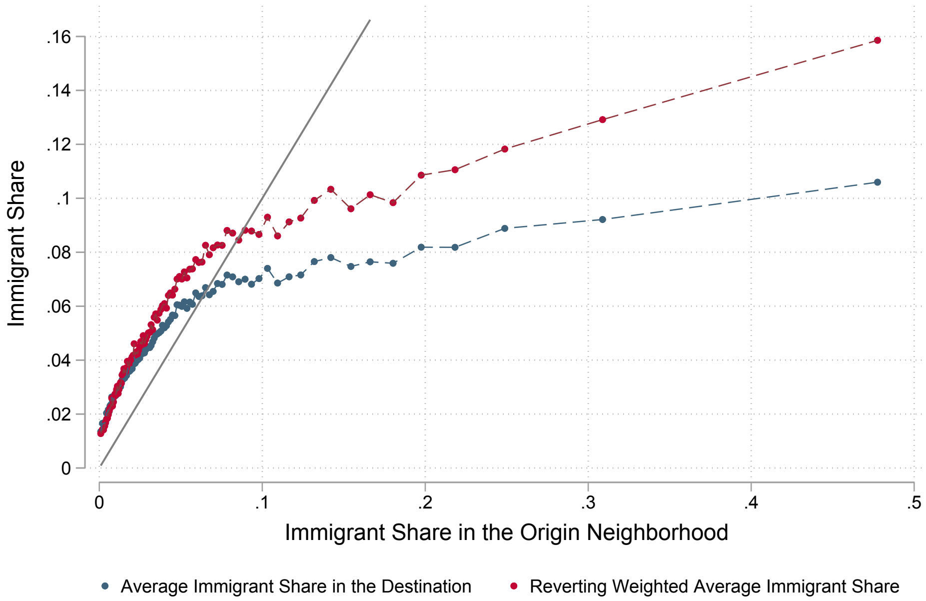


Figure 12. Association between origin's and destination's immigrant shares. This figure shows the relationship between the immigrant shares at the origin and destination for all Danish citizens and heads of households who moved to a different neighborhood. It plots the observed mean immigrant share at the destination (blue) and the mean Reverting Weighted Average Immigrant Share (red) against the mean share of immigrants in the origin neighborhood.

9 Conclusion

Immigration has become the most salient and divisive issue in the political discourse of most developed countries. For instance, 31 percent of EU citizens see it as a problem, and another 38 percent as both a "problem and an opportunity." Its labor market effects have been widely studied by economists, but nowadays seem to play a secondary role in the minds of voters. This could be due to the fact that other shocks to the labor market – trade, offshoring, robotization, skill-biased technologies – may have turned out to be more prominent.

In contrast, day-to-day coexistence issues seem to have made it to the forefront, with 47% of citizens in the EU believing that the integration of immigrants in their countries has not been successful.[33] Therefore, in our view, the study of ethnic tensions and native flight should also come to the forefront of research on this topic.

In the US, a firmly grounded literature in Economics studies ethnic tensions at the neighborhood level. Originally, this literature focused on the segregation of the African-American population as an expression of generalized white racism. However, it later ex-

[33] All opinion statistics in this paragraphs come from Eurobarometer (2022). With regard to integration, 37% believe it has been "not very successful" and an additional 10% "not at all successful".

panded to include the study of neighborhood configurations that arose from that legacy and broadened its scope to include immigration and the Hispanic population. Nevertheless, the depth and breadth of the current literature on immigrant segregation are far from commensurate with its critical importance.

Additionally, most of the literature on ethnic segregation relies on static equilibrium models, which are well-suited to describing minority shares in decennial data or explaining why some neighborhoods become fully segregated. However, we argue that classical models fail to accurately depict reality in dynamic environments with a constant turnover of majority and minority populations, heterogeneity in locational matching quality, or diverging tastes for homophily.

We began with the development of a dynamic model of residential mobility that distinguishes among three mechanisms: composition-neutral churning, replacement without displacement, and taste-based native flight. This model emphasizes the necessity of high-frequency individual or household panel data for identifying these mechanisms. Demographic change may arise through voluntary turnover, resulting in replacement without displacement. Introducing a native distaste for minority presence generates avoidance and flight, which increases immigrant concentration in equilibrium. When native attitudes are heterogeneous, segregationist outflows and neutral native inflows partially offset each other, yielding intermediate long-run minority shares. Overall, the model shows that low-frequency, aggregate data can obscure these dynamics and misrepresent the extent of native flight and the heterogeneity of the preferences that drive neighborhood change.

We then exploited a complete, georeferenced, individual-property panel dataset covering the entire population of Denmark from 1987 to 2017. During this period, there was a significant influx of non-Western immigrants, and in the later years, there was noticeable discontent among natives regarding their increasing numbers. Our findings conclusively demonstrate that many individuals expressed their dissatisfaction by voting with their feet and relocating away from neighborhoods where immigrants had settled.

The quality of the data allows us to conduct an exhaustive analysis and carefully assign causality. Unlike previous literature, we can model individual household decisions. This enables us to control for factors associated with high native turnover. We can also leverage the random nature of refugee placements to generate suitable quasi-experimental variation. In a demanding specification, we control for neighborhood-by-year fixed effects, which subsume local, time-specific shocks. Even then, we can identify flight of natives from buildings that hosted a higher proportion of non-Westerners, compared to similar buildings in the same area and year.

To the best of our knowledge, ours is the first study of tipping points in relation to individual "white" flight from a neighborhood, as opposed to the overall changes in its ethnic composition. This distinction is important because the departure of less tolerant residents may be offset by the arrival of more accommodating ones. Our study identifies an intriguing inverted U-shaped relationship between immigrant concentration and flight. As the immigrant population grows, the likelihood that native individuals will move out increases steadily, peaking at a foreign-born share of around 30%. However, beyond this threshold, the probability of relocation decreases, becoming similar to the baseline at around a 40% immigrant share. We interpret this as a sign of heterogeneous preferences: everyone who wanted to leave has already done so by that point.

Our study provides compelling evidence of heterogeneous behavior. Flight is more prevalent among older individuals and in multifamily buildings. Furthermore, our analysis shows that it only occurs in response to the presence of low-income immigrants. These findings underscore the importance of considering the nuanced effects of different immigrant profiles when examining the dynamics of native mobility.

Areas with a high immigrant population tend to attract new residents who are either non-Western immigrants or young, low-income Danes without children. Policymakers should pay attention to this sorting process of natives in minority neighborhoods, as the families moving in might not be the best equipped to manage potential inter-community conflict. Additionally, we find that families who move out tend to choose destinations with lower immigrant shares than predicted by a gravity model.

Our findings are relevant for policymakers and academics concerned with immigration and ethnic segregation. However, we acknowledge that their generalizability to other contexts remains uncertain. Nevertheless, our results demonstrate how social interactions influenced by native preferences can hinder neighborhood integration. It is also important to consider the composition of the native population in a given neighborhood. Consequently, conventional policy instruments that target specific geographic locations and focus solely on immigrant families may be ineffective at reducing the prevalence of segregation in the long run. To address concerns about the emergence of "parallel societies," policymakers should explore approaches that also target localized native behavior.

References

Accetturo, Antonio, Francesco Manaresi, Sauro Mocetti, and Elisabetta Olivieri, "Don't Stand so close to me: the urban impact of immigration," *Regional Science and Urban Economics*, 2014, *45*, 45–56.

Ahlfeldt, Gabriel M, C Soto Vieira, I Mulalic, and Daniel M Sturm, "The Geography of Life: Evidence From Copenhagen," *Working Paper*, 2025.

Amior, Michael, "The contribution of immigration to local labor market adjustment," *CEP Discussion Paper N.1678*, 2020.

_ , "Immigration, local crowd-out and undercoverage bias," *CEP Discussion Paper N.1669*, 2020.

Anderson, Michael L. and Jeremy Magruder, "Split-sample strategies for avoiding false discoveries," *NBER Working Paper n.23544*, July 2017.

Andersson, Henrik, Heléne Berg, and Matz Dahlberg, "Migrating natives and foreign immigration: Is there a preference for ethnic residential homogeneity?," *Journal of Urban Economics*, 2021, *121*, 103296.

_ **, Ina Blind, Fabian Brunåker, Matz Dahlberg, Greta Fredriksson, Jakob Granath, and Che-Yuan Liang**, "What's in a label? On neighbourhood labelling, stigma and housing prices," *Working Paper*, 2023.

Andries, Marianne, Leonardo Bursztyn, Thomas Chaney, and Milena Djourelova, "In their shoes," *Working Paper*, 2023.

Angrist, Joshua and Guido Imbens, "Identification and estimation of local average treatment effects," *NBER Working Paper n.0118*, 1995.

Ansala, Laura, Ulla Hämäläinen, and Matti Sarvimäki, "Age at arrival, parents and neighborhoods: understanding the educational attainment of immigrants' children," *Journal of Economic Geography*, 2020, *20* (2), 459–480.

Åslund, Olof and Peter Fredriksson, "Peer effects in welfare dependence: quasi-experimental evidence," *Journal of Human Resources*, 2009, *44* (3), 798–825.

_ **, John Östh, and Yves Zenou**, "How important is access to jobs? Old question—improved answer," *Journal of Economic Geography*, 2010, *10* (3), 389–422.

__ , **Per-Anders Edin, Peter Fredriksson, and Hans Grönqvist**, "Peers, neighborhoods, and immigrant student achievement: Evidence from a placement policy," *American Economic Journal: Applied Economics*, 2011, *3* (2), 67–95.

Azlor, Luz, Anna Piil Damm, and Marie L. Schultz-Nielsen, "Local Labour demand and immigrant employment," *Labour Economics*, 2020, *63.*

Bailey, Michael, Abhinav Gupta, Sebastian Hillenbrand, Theresa Kuchler, Robert Richmond, and Johannes Stroebel, "International trade and social connectedness," *Journal of International Economics*, 2021, *129*, 103418.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan, "A unified framework for measuring preferences for schools and neighborhoods," *Journal of Political Economy*, 2007, *115* (4), 588–638.

__ , __ , **and** __ , "A unified framework for measuring preferences for schools and neighborhoods," *Journal of political economy*, 2007, *115* (4), 588–638.

__ , **Hanming Fang, and Robert McMillan**, "Separate when equal? Racial inequality and residential segregation," *Journal of Urban Economics*, 2014, *82*, 32–48.

__ , **Marcus D Casey, W Ben McCartney, John Orellana-Li, and Calvin S Zhang**, "Distinguishing causes of neighborhood racial change: A nearest neighbor design," *American Economic Review*, 2025, *115* (11), 3999–4039.

Benjamini, Yoav, Abba M Krieger, and Daniel Yekutieli, "Adaptive linear step-up procedures that control the false discovery rate," *Biometrika*, 2006, *93* (3), 491–507.

Betts, Julian R and Robert W Fairlie, "Does immigration induce 'native flight'from public schools into private schools?," *Journal of Public Economics*, 2003, *87* (5-6), 987–1012.

Bjerre-Nielsen, Andreas and Mikkel Høst Gandil, "Attendance boundary policies and the limits to combating school segregation," *American Economic Journal: Economic Policy*, February 2024, *16* (1), 190–227.

Blair, Peter Q, "Beyond racial attitudes: The role of outside options in the dynamics of white flight," *NBER Working Paper n.31136*, 2023.

Bluestone, Daniel, "Framing landscape while building density: Chicago Courtyard Apartments, 1891–1929," *Journal of the Society of Architectural Historians*, 2017, *76* (4), 506–531.

Bond, Eric W. and N.Edward Coulson, "Externalities, filtering, and neighborhood change," *Journal of Urban Economics*, 1989, *26* (2), 231–249.

Borjas, George J., "Ethnicity, neighborhoods, and human-capital externalities," *The American Economic Review*, 1995, *85* (3), 365–390.

_ **and Joan Monras**, "The labour market consequences of refugee supply shocks," *Economic Policy*, 08 2017, *32* (91), 361–413.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "Quasi-experimental shift-share research designs," *The Review of Economic Studies*, 06 2021, *89* (1), 181–213.

Boustan, Leah Platt, "Was postwar suburbanization "white flight"? Evidence from the black migration," *The Quarterly Journal of Economics*, 2010, *125* (1), 417–443.

Bursztyn, Leonardo, Thomas Chaney, Tarek A. Hassan, and Aakaash Rao, "The immigrant next door," *American Economic Review*, February 2024, *114* (2), 348–84.

Böhlmark, Anders and Alexander Willén, "Tipping and the effects of segregation," *American Economic Journal: Applied Economics*, January 2020, *12* (1), 318–47.

Card, David, "The impact of the Mariel boatlift on the Miami labor market," *Ilr Review*, 1990, *43* (2), 245–257.

_ , "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 2001, *19* (1), 22–64.

_ **, Alexandre Mas, and Jesse Rothstein**, "Tipping and the dynamics of segregation," *The Quarterly Journal of Economics*, 2008, *123* (1), 177–218.

Cascio, Elizabeth U and Ethan G Lewis, "Cracks in the melting pot: immigration, school choice, and segregation," *American Economic Journal: Economic Policy*, 2012, *4* (3), 91–117.

Cutler, David M. and Edward L. Glaeser, "Are ghettos good or bad?," *The Quarterly Journal of Economics*, 1997, *112* (3), 827–872.

Cutler, David M, Edward L Glaeser, and Jacob L Vigdor, "The rise and decline of the American ghetto," *Journal of Political Economy*, 1999, *107* (3), 455–506.

_ **, Edward L. Glaeser, and Jacob L Vigdor**, "Is the melting pot still hot? Explaining the resurgence of immigrant segregation," *The Review of Economics and Statistics*, 2008, *90* (3), 478–497.

Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor, "When are ghettos bad? Lessons from immigrant segregation in the United States," *Journal of Urban Economics*, 2008, *63* (3), 759–774.

Damm, Anna Piil, "Determinants of recent immigrants' location choices: quasi-experimental evidence," *Journal of Population Economics*, 2009, *22*, 145–174.

_ , "Neighborhood quality and labor market outcomes: Evidence from quasi-random neighborhood assignment of immigrants," *Journal of Urban Economics*, 2014, *79*, 139–166.

_ **, Ahmad Hassani, and Jonas Sørensen**, "Place-Based Policies in Deprived Neighbourhoods: Opportunities for Preexisting Residents and Neighbourhood Revitalisation?," *IZA Discussion Paper N.17843*, 2025.

_ **and Christian Dustmann**, "Does growing up in a high crime neighborhood affect youth criminal behavior?," *American Economic Review*, June 2014, *104* (6), 1806–32.

_ **and Marie L. Schultz-Nielsen**, "Danish neighbourhoods: construction and relevance for measurement of residential segregation," *Nationaløkonomisk tidsskrift*, 2008, *146* (3), 241–262.

Davis, Donald R., Matthew Easton, and Stephan Thies, "Segregation, Spillovers, and the Locus of Racial Change," NBER Working Papers 34246, National Bureau of Economic Research, Inc Sep 2025.

Derenoncourt, Ellora, "Can you move to opportunity? Evidence from the Great Migration," *American Economic Review*, 2022, *112*, 145–174.

Dustmann, Christian, Uta Schönberg, and Jan Stuhler, "Labor supply shocks, native wages, and the adjustment of local employment," *The Quarterly Journal of Economics*, 2017, *132* (1), 435–483.

Edin, Per-Anders, Peter Fredriksson, and Olof Åslund, "Ethnic enclaves and the economic success of immigrants—Evidence from a natural experiment," *The Quarterly Journal of Economics*, 2003, *118* (1), 329–357.

Egge, Marit and Randi Solhjell, *Parallellsamfunn: En del av den norske virkeligheten?*, Politihøgskolen, 2018.

Eurobarometer, "Integration of immigrants in the European Union," *Special Eurobarometer Report*, 2022, *519.*

Farre, Lidia, Francesc Ortega, and Ryuichi Tanaka, "Immigration and the public–private school choice," *Labour Economics*, 2018, *51*, 184–201.

Foged, Mette and Giovanni Peri, "Immigrants' effect on native workers: New analysis on longitudinal data," *American Economic Journal: Applied Economics*, 2016, *8* (2), 1–34.

Frankel, David M. and Ady Pauzner, "Expectations and the timing of neighborhood change," *Journal of Urban Economics*, March 2002, *51* (2), 295–314.

Glaeser, Edward L., Jed Kolko, and Albert Saiz, "Consumer city," *Journal of Economic Geography*, 2001, *1* (1), 27–50.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik instruments: What, when, why, and how," *American Economic Review*, 2020, *110* (8), 2586–2624.

Govind, Yajna, Jack Melbourne, Sara Signorelli, and Edith Zink, "The Making of a Ghetto Place-Based Policies, Labeling, and Impacts on Neighborhoods and Individuals," *RFBerlin Discussion Paper No. 83/25*, 2025.

Graaff, Thomas De, Michiel Van Leuvensteijn, and Casper Van Ewijk, "Homeownership, social renting and labor mobility accross Europe," *Homeownership and the labour market in Europe. Oxford University Press, Oxford*, 2009, pp. 53–81.

Gutierrez-i-Puigarnau, Eva, Ismir Mulalic, and Jos N. van Ommeren, "Do rich households live farther away from their workplaces?," *Journal of Economic Geography*, 2016, *16*, 177–201.

Han, Joseph, Jinwook Hur, Jongkwan Lee, and Hyunjoo Yang, "To move or not to move? Immigration and natives' neighborhood choices in Seoul, Korea," *Journal of Economic Geography*, 2022, *22* (4), 779–799.

Hise, Greg, "From Roadside Camps to Garden Homes: Housing and Community Planning for California's Migrant Work Force, 1935-1941," *Perspectives in Vernacular Architecture*, 1995, *5*, 243–258.

Jaeger, David A., Joakim Ruist, and Jan Stuhler, "Shift-share Instruments and the Impact of Immigration," *NBER Working Paper 24285*, 2018.

Jensen, Peter and Peder J Pedersen, "To stay or not to stay? Out-migration of immigrants from Denmark," *International Migration*, 2007, *45* (5), 87–113.

Jia, Ning, Raven Molloy, Christopher Smith, and Abigail Wozniak, "The economics of internal migration: Advances and policy questions," *Journal of Economic Literature*, 2023, *61* (1), 144–180.

Kain, John F, "Housing segregation, negro employment, and metropolitan decentralization," *The Quarterly Journal of Economics*, 1968, *82* (2), 175–197.

King, A Thomas and Peter Mieszkowski, "Racial discrimination, segregation, and the price of housing," *Journal of Political Economy*, 1973, *81* (3), 590–606.

Massey, Douglas S, "American apartheid: Segregation and the making of the underclass," *American Journal of Sociology*, 1990, *96* (2), 329–357.

McFadden, Daniel, "Economic choices," *American economic review*, 2001, *91* (3), 351–378.

Melitz, Jacques and Farid Toubal, "Native language, spoken language, translation and trade," *Journal of International Economics*, 2014, *93* (2), 351–363.

Miyao, Takahiro, "Dynamic instability of a mixed city in the presence of neighborhood externalities," *The American Economic Review*, 1978, *68* (3), 454–463.

_ , "Dynamic stability of an open city with many household classes," *Journal of Urban Economics*, 1979, *6* (3), 292–298.

Molla, Hemrin, Caroline Rhawi, and Elina Lampi, "Name matters! The cost of having a foreign-sounding name in the Swedish private housing market," *Plos One*, 2022, *17* (6), e0268840.

Monras, Joan, "Local adjustment to immigrant-driven labor supply shocks," *Journal of Human Capital*, 2021, *15* (1), 204–235.

Moraga, Jesús FH, Ada Ferrer-i-Carbonell, and Albert Saiz, "Immigrant locations and native residential preferences: Emerging ghettos or new communities?," *Journal of Urban Economics*, 2019, *112*, 133–151.

Mulalic, Ismir and Jan Rouwendal, "Does improving public transport decrease car ownership? Evidence from a residential sorting model for the Copenhagen metropolitan area," *Regional Science and Urban Economics*, 2020, *83*, 103543.

Munch, Jakob Roland and Michael Svarer, "Rent control and tenancy duration," *Journal of Urban Economics*, 2002, *52* (3), 542–560.

Ortega, Javier and Gregory Verdugo, "Who stays and who leaves? Immigration and the selection of natives across locations," *Journal of Economic Geography*, 2022, *22* (2), 221–260.

Pennington, Kate, "Does building new housing cause displacement?: The supply and demand effects of construction in San Francisco," *Working Paper*, 2021.

Pew Research Center, "Religious composition by country, 2010-2050," https://www.pewresearch.org/religion/2015/04/02/religious-projection-table/, 2015.

Putnam, Robert D, *Bowling alone: The collapse and revival of American community*, Simon and schuster, 2000.

Romano, Joseph P and Michael Wolf, "Exact and approximate stepdown methods for multiple hypothesis testing," *Journal of the American Statistical Association*, 2005, *100* (469), 94–108.

Sá, Filipa, "Immigration and house prices in the UK," *The Economic Journal*, 2015, *125* (587), 1393–1424.

Saiz, Albert, "Room in the kitchen for the melting pot: Immigration and rental prices," *Review of Economics and Statistics*, 2003, *85* (3), 502–521.

_ , "Immigration and housing rents in American cities," *Journal of Urban Economics*, 2007, *61* (2), 345–371.

_ **and Susan Wachter**, "Immigration and the neighborhood," *American Economic Journal: Economic Policy*, May 2011, *3* (2), 169–88.

Salazar, Arianna, "The shape of segregation: The role of urban form in immigrant assimilation," *Cities*, 2020, *106*, 102852.

Samuelson, Paul A, "What classical and neoclassical monetary theory really was," *The Canadian Journal of Economics/Revue canadienne d'Economique*, 1968, *1* (1), 1–15.

Schelling, Thomas C., "Dynamic models of segregation," *The Journal of Mathematical Sociology*, 1971, *1* (2), 143–186.

Shertzer, Allison and Randall P Walsh, "Racial sorting and the emergence of segregation in American cities," *Review of Economics and Statistics*, 2019, *101* (3), 415–427.

Silver, Laura and Courtney Johnson, "Younger adults in Western Europe are more socially, politically progressive than older age groups," Pew Research Center, Short Reads. https://www.pewresearch.org/short-reads/2018/10/30/younger-western-europeans-more-socially-and-politically-progressive/#:~:text=Younger%20Europeans%20in%20most%20countries,of%20those%2050%20and%20older, 2018.

Statistics Denmark, "Danish Industrial Classification of All Economic Activities 1993," 1995.

_ , "Dansk Branchekode 2007," 2007.

Taeuber, Karl E and Alma F Taeuber, *Residential segregation and neighborhood change*, Transaction Publishers, 1965.

Verdugo, Gregory, "Public housing magnets: public housing supply and immigrants' location choices," *Journal of Economic Geography*, 2016, *16* (1), 237–265.

_ **and Sorana Toma**, "Can public housing decrease segregation? Lessons and challenges from non-European immigration in France: Can public housing decrease segregation?," *Demography*, 2018, *55* (5), 1803–1828.

Waldfogel, Joel, "The median voter and the median consumer: Local private goods and population composition," *Journal of Urban Economics*, 2008, *63* (2), 567–582.

White, Justin S, Rita Hamad, Xinjun Li, Sanjay Basu, Henrik Ohlsson, Jan Sundquist, and Kristina Sundquist, "Long-term effects of neighbourhood deprivation on diabetes risk: quasi-experimental evidence from a refugee dispersal policy in Sweden," *The Lancet Diabetes & Endocrinology*, 2016, *4* (6), 517–524.

World Bank, "World Bank Country and Lending Groups," https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups, 2022.

Yinger, John, "Racial prejudice and racial residential segregation in an urban model," *Journal of Urban Economics*, 1976, *3* (4), 383–396.

Online Appendix for "Immigrants and Native Flight: Geographic Extent and Heterogeneous Preferences"

Bence Boje-Kovacs
Aalborg University

Ismir Mulalic
CBS

Albert Saiz
MIT

Vinicios Sant'Anna
California Polytechnic State University

Marie Louise Schultz-Nielsen
ROCKWOOL Foundation Research

January 2026

Table of Contents

A A Dynamic Theory of Immigration and Residential Mobility

The economy consists of two locations: neighborhood A and the rest of the world, denoted by W. Each resident lives in a housing unit in either A or W in every period t. Assume that W has a perfectly elastic supply, while neighborhood A has a fixed maximum capacity $\overline{H}^A$. Let the measure of native residents in the economy be one. We can write the allocation of residents between A and W in period t as follows:

$$1 = H_t^A + H_t^W, \textit{with } H_t^A \leq \bar{H}^A < H_t^W \tag{A.1}$$

where H_t^A and H_t^W represent the measure (share) of housing units (residents) in neighborhood A and in the rest of the world in period t, respectively.

Baseline Preferences Assume that each resident i in period t obtains an indirect utility from living in neighborhood A or the rest of the world, W, in the form:

$$\begin{aligned} V_{i,t}^A &= \varepsilon_{i,t} - \Omega_t, \\ V_{i,t}^W &= 0, \quad \forall i \end{aligned} \tag{A.2}$$

where $\varepsilon_{i,t} > 0$ is an idiosyncratic preference parameter for neighborhood A drawn from an exponential distribution with parameter λ, i.i.d. across residents and time.[1] Ω_t is the endogenous occupancy cost paid by residents to live in A in period t. It includes the rents and other non-monetary costs associated with finding and living in a unit in A, such as searching or waiting for a vacancy on the market. Residents living in W get a fixed reservation utility normalized to zero, and their occupancy cost is normalized to zero, $\Omega^W = 0$.[2]

At the beginning of every period t, each resident i draws their new idiosyncratic preference parameter, $\varepsilon_{i,t}$, and may decide to move. A short-term equilibrium then takes

[1]While the literature on residential mobility typically uses Extreme Value distributions (e.g., Type I Extreme Value or Fréchet), we employ the exponential distribution for the idiosyncratic preference parameter for three reasons. First, it naturally captures a structure where most residents have a modest willingness to pay for neighborhood A, with a declining probability of higher valuations. Second, the exponential distribution has a thin upper tail, which implies low probabilities of extremely high preference draws for A, making it more realistic that having only a small fraction of residents would be willing to pay very high premiums for the amenities of neighborhood A. Third, it yields tractable closed-form solutions for the equilibrium cost of occupancy and mobility flows.

[2]While we could model preferences for W with another idiosyncratic preference term and associated cost of occupancy, the normalization allows us to significantly improve model tractability without loss of generality. Future research could explore a version of this model with multiple neighborhoods, leveraging more complex structures for these preferences.

the form of a decision rule by which, in each period, residents will decide whether to stay in their neighborhood or move to the other based on the cost of occupancy in neighborhood A and the draws of the idiosyncratic preference.

Scenario 1. No Moving Costs

We begin by assuming that residents can sort between the two locations without incurring any moving costs. Resident i will choose to live in A if:

$$\varepsilon_{i,t} - \Omega_t \ > \ 0 \quad \Longleftrightarrow \quad \varepsilon_{i,t} \ > \ \Omega_t \tag{A.3}$$

In an equilibrium where the maximum capacity of A is binding, exactly $\bar{H}^A$ residents must reside in A. The share of residents whose idiosyncratic draw $\varepsilon_{i,t}$ exceeds Ω_t must therefore be $\bar{H}^A$. For an exponential distribution with parameter λ, this can be represented by the following condition:[3]

$$e^{-\lambda\Omega_t} \ = \ \bar{H}^A \tag{A.4}$$

Solving for Ω_t yields the equilibrium occupancy cost (Ω_t^*) in neighborhood A, under no moving costs:

$$\boxed{\Omega_t^* \ = \ -\frac{1}{\lambda}\ln\left(\bar{H}^A\right)} \tag{A.5}$$

Intuitively, as the housing capacity of A ($\bar{H}^A$) grows, the equilibrium occupancy cost Ω_t^* must fall so that a larger share of the population finds A more attractive than W. A larger λ implies a thinner upper tail on ε, making it less likely for residents to have extreme preferences for A, consequently reducing Ω_t^*.

Scenario 2. Introducing Moving Costs

Now, we add moving costs to the baseline scenario. If the resident moves from A to W, she incurs a cost $C \geq 0$, which is the same for all residents, for all periods, and is publicly known.[4] To ensure an equilibrium with positive churn, we also assume that moving costs are small and do not exceed the occupancy cost ($C \leq \Omega_t$).[5]

[3]Notice that $\mathbb{P}\left(\varepsilon_{i,t} \geq \Omega_t\right) \ = \ e^{-\lambda\Omega_t}$.

[4]Introducing moving costs for relocation from W to A does not qualitatively affect the model's results. It lowers the equilibrium occupancy cost ($C \leq \Omega_t^*$) and churn, but the capacity constraint still binds. The long-run immigrant share is unchanged (with slower convergence if natives are indifferent), and all qualitative results hold when natives dislike immigrant concentration. Overall, introducing moving costs for relocation from W to A mainly dampens adjustment speed via lower churn.

[5]This condition is necessary to ensure an equilibrium with positive moving rates. If $C > \Omega_t$, no resident would move out of A. Because the exponential distribution is supported on the interval $[0, \infty)$,

Moves from W to A. A resident in W at the beginning of period t compares the payoff from staying in W with the one from moving $(\varepsilon_{i,t} - \Omega_t)$. They move to A if

$$\varepsilon_{i,t} - \Omega_t \; > \; 0 \quad \Longleftrightarrow \quad \varepsilon_{i,t} \; > \; \Omega_t \tag{A.6}$$

Denote by $M_t^{W\to A}$ the measure of residents that moved from W to A, and $M_t^{W\to W}$ the measure that stayed in W. Under the scenario where the maximum capacity of A is binding, $(H_{t-1}^A = \bar{H}^A)$, we can write the measure of residents moving from W to A as:

$$M_t^{W\to A} \; = \; H_{t-1}^W \cdot e^{-\lambda\,\Omega_t} \; = \; (1 - \bar{H}^A) \cdot e^{-\lambda\,\Omega_t}, \tag{A.7}$$

and the corresponding measure staying in W as:

$$M_t^{W\to W} \; = \; (1 - \bar{H}^A) \cdot (1 - e^{-\lambda\,\Omega_t}) \tag{A.8}$$

Moves from A to W. Similarly, a resident in A at the beginning of period t decides whether to remain or move. If she stays in A, she collects the payoff $(\varepsilon_{i,t} - \Omega_t)$, given by the new idiosyncratic utility draw for A minus the occupancy cost. If she moves to W, she pays the moving cost (C) and obtains the zero payoff from living in W.

Hence, people in A will move to W if:

$$-C \; \geq \; \varepsilon_{i,t} - \Omega_t \quad \Longleftrightarrow \quad \varepsilon_{i,t} \; \leq \; \Omega_t - C. \tag{A.9}$$

Under the scenario where the maximum capacity of A is binding, $(H_{t-1}^A = \bar{H}^A)$, we can write the measure of people moving out of A $(M_t^{A\to W})$ and the measure staying in A $(M_t^{A\to A})$ as follows:

$$M_t^{A\to W} \; = \; \bar{H}^A \cdot \left(1 - e^{-\lambda\left(\Omega_t - C\right)}\right), \tag{A.10}$$

$$M_t^{A\to A} \; = \; \bar{H}^A \cdot e^{-\lambda\left(\Omega_t - C\right)} \tag{A.11}$$

Because the population is normalized to one, these measures of migration are equivalent to their population shares. Therefore, we can also refer to the terms $M_t^{A\to W}$ and $M_t^{W\to A}$ as the residential *churn* rates Hence, Equations (A.7) and (A.10) provide important insights about residential churn within our framework.

Result A.1 ***(Churn and the Location Matching Parameter)*** *The distribution parameter λ, which shapes the spread of the idiosyncratic draws $\varepsilon_{i,t}$ has a positive re-*

there are no negative preference shocks that could compensate for excessively high moving costs relative to the cost of occupancy. In such a scenario, residents would be locked into A, and the condition $\varepsilon_{i,t} < \Omega_t - C$ would require draws outside the distribution's support.

lationship with churn. A larger λ implies a thinner upper tail, making it less likely for residents to have high idiosyncratic preference draws for A, thus generally increasing *outflows from A to W, and* decreasing *flows from W to A:*

$$
\begin{aligned}
\frac{\partial M^{A\to W}}{\partial \lambda} &= (\Omega_t - C)\overline{H}^A e^{-\lambda(\Omega_t - C)} > 0, \quad with \quad (\Omega_t > C) \\
\frac{\partial M^{W\to A}}{\partial \lambda} &= -\Omega_t(1-\overline{H}^A)e^{-\lambda\Omega_t} < 0
\end{aligned}
\tag{A.12}
$$

Result A.2 ***(Churn and Moving Costs)*** *The moving cost C displays a negative relationship with respect to neighborhood's A churn rate. A higher C, discourages residential mobility from A to W, thereby lowering the churn rate:*

$$
\frac{\partial M^{A\to W}}{\partial C} = -\lambda\overline{H}^A e^{-\lambda(\Omega_t - C)} < 0 \tag{A.13}
$$

Dynamic Equilibrium Condition Because neighborhood A can accommodate at most $\overline{H}^A$ of the population, the new fraction of people living in A at the end of period t must satisfy the housing availability constraint:

$$
H_t^A \;=\; M_t^{A\to A} \;+\; M_t^{W\to A} \;\leq\; \overline{H}^A \tag{A.14}
$$

In words, the measure of inhabitants in A must be equal to the sum of those that were already living in A ($M_t^{A\to A}$) and those that recently moved to A ($M_t^{W\to A}$). In a *binding* equilibrium (where A is always exactly filled to capacity), we have:

$$
\overline{H}^A \;=\; \left[\overline{H}^A \cdot e^{-\lambda(\Omega_t - C)}\right] \;+\; \left[\left(1-\overline{H}^A\right) \cdot e^{-\lambda\Omega_t}\right] \tag{A.15}
$$

Hence, solving for the *equilibrium occupancy cost* Ω^*:

$$
\boxed{\Omega^* \;=\; -\frac{1}{\lambda}\ln\left[\frac{\overline{H}^A}{1+(e^{\lambda C}-1)\overline{H}^A}\right]} \tag{A.16}
$$

The equilibrium occupancy cost Ω^* must adjust so that the sum of those staying in A and those moving to A precisely fills all available housing in A, but does not exceed the maximum capacity $\overline{H}^A$. Alternatively, in equilibrium, the mass of movers across the two locations must be the same: $M^{A\to W} = M^{W\to A}$. Notice that, under positive moving costs, the occupancy cost must be higher than in the baseline without them. As some residents are "sticky" and will not leave, individual with high idiosyncratic positive shocks need to bid more aggressively to get in to neighborhood A. Notice also that the equation nests to the baseline equilibrium when moving costs are zero.

Scenario 3. Introducing Immigration

Now, let's introduce immigrants to the model. Assume that a measure ϕ of immigrants arrives from abroad to an economy under the equilibrium defined in Scenario 2.[6] Therefore, the new measure of residents increases by ϕ:

$$1 + \phi = \overline{H}^A + H_t^W \tag{A.17}$$

Because the housing capacity in A is constrained, the arrival of more residents necessarily implies the expansion of the housing stock in the rest of the world, W, to absorb the new residents. Let's start by assuming that immigrants are identical to natives. So their preferences are given by Equation (A.2). This implies that, at any period, there will be a potentially higher mass of people willing to move from W to A. Thus, the new measure of residents moving from W to A is

$$M_t^{W\to A} \;=\; H_{t-1}^W \cdot e^{-\lambda\,\Omega_t} \;=\; (1+\phi-\overline{H}^A)\cdot e^{-\lambda\,\Omega_t} \tag{A.18}$$

Solving for the new occupancy cost gives us:

$$\boxed{\Omega^* \;=\; -\frac{1}{\lambda}\ln\left[\frac{\overline{H}^A}{1+\phi+(e^{\lambda C}-1)\overline{H}^A}\right]} \tag{A.19}$$

Result A.3 ***(Immigration and the Cost of Occupancy)*** *More immigrants, in this case, simply imply more residents, resulting in higher occupancy costs. With more residents, we have an expected larger mass of people drawing high values of ε, increasing housing demand for A, and consequently, the cost of occupancy in there* (Ω^*).

The dynamics of immigrant presence The model predicts a specific pattern for the distribution of immigrants between neighborhoods. Let's denote the number of immigrants living in neighborhood A at time t as ϕ_t^A and those living in W as $\phi_t^W = \phi - \phi_t^A$. When immigrants first arrive ($t = 0$), we assume that they all initially settle in W ($\phi_0^A = 0$). We also express immigrant concentration in A as $x_t^A = \frac{\phi_t^A}{\overline{H}^A}$ (the *share* of

[6] We also assume that the mass of immigrant arrivals exceeds the size of neighborhood A, $\overline{H}^A \leq \phi$. This is a more realistic assumption, which implies a non-trivial measure of immigrant arrivals and allows for a scenario with complete immigrant in A, like in the tipping point theories (Schelling, 1971; Card et al., 2008).

residents in A who are immigrants), and $x_t^W = \frac{\phi_t^W}{H^W}$ (the analogous *share* of residents in W who are immigrants) at time t.

Since immigrants and natives have identical preference distributions, these measures of immigrant concentration will also determine the proportions of immigrants moving in and out of neighborhood A. The evolution of immigrant presence in A can be written as:

$$\Delta\phi_t^A = M^{W\to A} \cdot x_{t-1}^W - M^{A\to W} \cdot x_{t-1}^A \tag{A.20}$$

where $\Delta\phi_t^A$ is the change in the number of immigrants living in A at time t. The first term is the inflow of immigrants moving into A (arrivers): the product of the mass of residents moving from W to A and the immigrant share in W at $t-1$ (the probability that an arriver is an immigrant). The second term is the outflow of immigrants from A (movers): the product of the mass of residents moving from A to W and the immigrant share in A at $t-1$ (the probability that a mover is an immigrant).

Result A.4 ***(Inflow of Immigrants to A)*** *Each year, the churn measures in and out of A must be the same ($M^{A\to W} = M^{W\to A}$), which implies that immigrant presence in A increases only when their concentration in W exceeds their concentration in A:*

$$\Delta\phi_t^A > 0 \iff x_{t-1}^W > x_{t-1}^A \tag{A.21}$$

Initially, all immigrants are in W, so $\phi_0^W = \phi$, implying positive net immigrant flows to A in $t = 1$, $\Delta\phi_1^A > 0$. Over time, as residents churn due to new preference shocks, immigrants gradually move into neighborhood A, following Equation (A.20), as long as the presence of immigrants is unbalanced in favor of W. Hence, the model explains how demographic composition changes can occur naturally through the housing market's ordinary churn, without requiring changes in individual preferences or behavior.

Result A.5 ***(Dynamic Equilibrium Share of Immigrants)*** *From Equation (A.20), we can identify the time when neighborhood A reaches a stable distribution of immigrants across space. Its expected change in immigrant shares becomes zero when:*

$$\Delta\phi_t^A = 0 \iff x_{t-1}^W = x_{t-1}^A \iff x^* = \frac{\phi}{1+\phi}. \tag{A.22}$$

This result demonstrates that, in the absence of preference differences between immigrants and natives, residential churn ultimately leads to proportional representation of immigrants across neighborhoods. At this point, the probability that a vacated unit in A is filled by an immigrant exactly matches the probability that the unit is vacated by an immigrant, resulting in a dynamic equilibrium of the local demographic composition.

This result resembles the Ergodic Hypothesis for income distribution as established by Samuelson (1968). Regardless of how we initially distribute immigrants and natives between neighborhoods, the system eventually converges to a unique ergodic state after a sufficiently long time. Therefore, we refer to the long-run stable share of immigrants in A, x^*, as the ergodic share.

Result A.6 ***(Moving Costs, Residential churn, and the Speed of Demographic Transition)*** *Neighborhoods with low moving costs, and consequently high churn, experience a faster demographic transition to the ergodic equilibrium.*

From Result (A.2) and Equation (A.13), we show the negative relationship between moving costs and churn. This reveals that lower moving costs (C), lead to higher churn rates ($M^{A \to W}$), and lower costs of occupancy. Equation (A.20) further demonstrates that lower moving costs accelerate immigrant inflows to A during the transition stage.

Moving costs (and, consequently, the baseline churn rate of a neighborhood) function as a regulator of the speed of the demographic transition. For a given distribution of immigrants across neighborhoods, higher churn rates produce larger changes in immigrant arrivals per period, hastening convergence to long-run equilibrium. This result implies that, even when their preferences are identical to those of natives, immigrants concentrate more rapidly in neighborhoods with lower occupancy costs (e.g., lower rents or shorter wait times) and higher residential turnover.

Result A.7 ***(Residential churn and the Immigrant Concentration)*** *Areas with higher baseline churn exhibit systematically higher immigrant presence throughout the transition period.*

This result has important implications for empirical analyses. Areas with different moving costs—such as rental market regulations or the prevalence of social housing—will experience different rates of demographic transition following immigration shocks. Failing to properly account for differences in baseline churn across neighborhoods can lead researchers to draw misleading conclusions about underlying preference-based sorting.

Figure A.1 illustrates this point. It plots the evolution of the immigrant share in neighborhood A (x_t^A) under three scenarios for the moving cost C: low (light blue), medium (intermediate blue), and high (dark blue), while holding all other parameters constant. Starting from the initial condition in which all immigrants reside in W, the figure illustrates the convergence to the ergodic equilibrium share (x^*). As expected, higher moving costs substantially slow the rate at which the immigrant share converges to its long-run ergodic equilibrium level. At t = 5, for example, the immigrant share

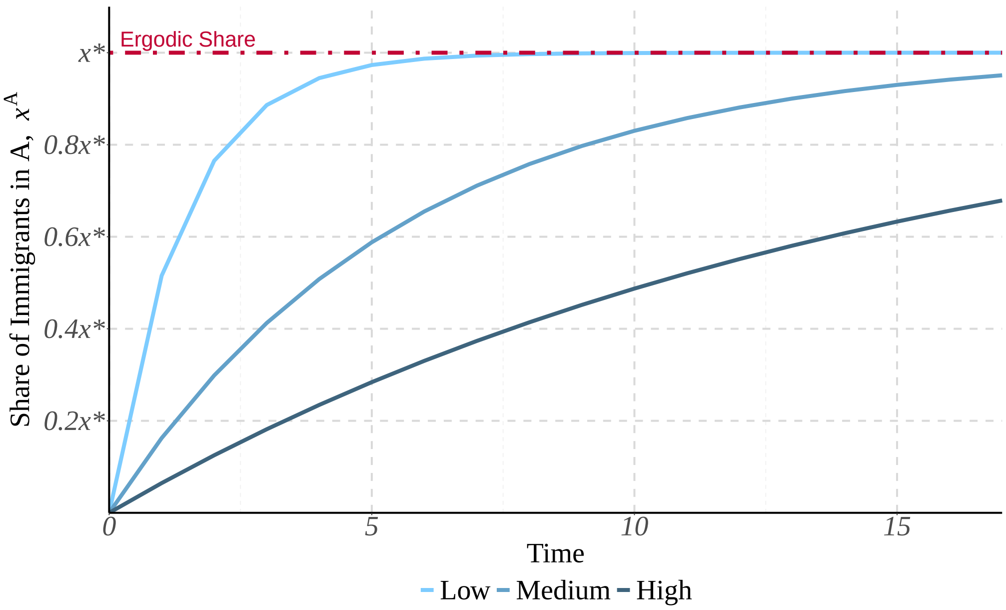


Figure A.1. Immigrant Share Evolution and Moving Costs. This figure illustrates the evolution of immigrant share in A over time for different values of moving cost, C: low (light blue), medium (intermediate blue), and high (dark blue).

reaches approximately 95% of the ergodic share under low moving costs, 60% under medium costs, and 30% under high costs. In a scenario with a high baseline churn rate, researchers may mistake the rapid adjustment for the presence of homophilic preferences (e.g., white flight).

Result A.8 ***(Replacement Without Displacement)*** *The immigrant share in Neighborhood A increases through a natural churn process. Natives who draw low local utility shocks voluntarily leave, creating vacancies that are disproportionately filled by immigrants. This mechanism increases the immigrant share without directly displacing natives.*

This result highlights the importance of considering individual-level decisions rather than relying solely on neighborhood-level compositional changes when studying residential sorting and segregation.

Scenario 4. Introducing Ethnic Preferences

We now extend our model to incorporate heterogeneous preferences for homophily. Immigrants maintain the same preferences as before, while natives experience a utility *penalty* with respect to the fraction of immigrants in neighborhood A. The indirect utility functions for natives in period t are now:

$$\begin{aligned} V_{i,t}^{A}(N) &= \varepsilon_{i,t} - \Omega_t - \delta \cdot ln(1 + x_t^A), \\ V_{i,t}^{W}(N) &= 0, \quad \forall i \end{aligned} \tag{A.23}$$

where $\delta > 0$ is a parameter that modulates the distaste of natives for immigrant concentration in neighborhood A. x_t^A.[7] For natives in W to move to A we now need that:

$$\varepsilon_{i,t} - \Omega_t - \delta \cdot ln(1 + x_t^A) > 0 \Leftrightarrow \varepsilon_{i,t} > \Omega_t + \delta \cdot ln(1 + x_t^A) \tag{A.24}$$

Therefore, measure of natives moving from W to A is:

$$M_t^{W \to A}(N) = (H_{t-1}^W - \phi_{t-1}^W) \cdot e^{-\lambda(\Omega_t + \delta \cdot ln(1 + x_t^A))} \tag{A.25}$$

Result A.9 ***(Distaste and Native Avoidance)*** *From Equation (A.25), the larger the distaste parameter δ or the share of immigrants x^A, the fewer natives will move to A, a pattern we can define as* native avoidance*:*

$$\frac{\partial M_t^{W \to A}(N)}{\partial \delta} < 0 \qquad \& \qquad \frac{\partial M_t^{W \to A}(N)}{\partial x^A} < 0. \tag{A.26}$$

For immigrants, the moving condition remains the same as before:

$$M_t^{W \to A}(I) = \phi_{t-1}^W \cdot e^{-\lambda \Omega_t} \tag{A.27}$$

Combining both groups, the total moves from W to A at any period are at:

$$\begin{aligned} M_t^{W \to A} =& M_t^{W \to A}(N) + M_t^{W \to A}(I) \\ =& (H_{t-1}^W - \phi_{t-1}^W) \cdot e^{-\lambda \Omega_t - \lambda \delta \cdot ln(1 + x_t^A)} + \phi_{t-1}^W \cdot e^{-\lambda \Omega_t} \end{aligned} \tag{A.28}$$

Conversely, the measure of natives moving from A to W is now:

$$M_t^{A \to W}(N) = (\overline{H}^A - \phi_{t-1}^A) \cdot \left[1 - e^{-\lambda(\Omega_t - C + \delta \cdot ln(1 + x_t^A))}\right] \tag{A.29}$$

Result A.10 ***(Distaste and Native Flight)*** *From Equation (A.29) we find that the larger the distaste parameter δ or the share of immigrants x^A, the larger the measure of*

[7] Replacing the log penalty in Equation (A.23) with a linear ($\delta \cdot x$) or exponential ($\delta \cdot e^{1+x}$) functional form leaves the qualitative results unchanged, while only rescaling magnitudes: the steepens of the utility penalty from xenophobia increases (log < linear < exponential), implying (for fixed δ) progressively higher x^* and slower convergence.

natives that will move out of A, $M_t^{A\rightarrow W}(N)$:

$$\frac{\partial M_t^{A\rightarrow W}(N)}{\partial \delta} > 0 \qquad \& \qquad \frac{\partial M_t^{A\rightarrow W}(N)}{\partial x^A} > 0. \tag{A.30}$$

For immigrants, the condition fro moving out of A remains unchanged:

$$M_t^{A\rightarrow W}(I) = \phi_{t-1}^A \cdot \left[1 - e^{-\lambda(\Omega_t - C)}\right] \tag{A.31}$$

Therefore, total moves from A to W combine both groups':

$$\begin{aligned} M_t^{A\rightarrow W} =& M_t^{A\rightarrow W}(N) + M_t^{A\rightarrow W}(I) \\ =& (\overline{H}^A - \phi_{t-1}^A) \cdot \left[1 - e^{-\lambda(\Omega_t - C + \delta \cdot ln(1+x_t^A))}\right] + \phi_{t-1}^A \cdot \left[1 - e^{-\lambda(\Omega_t - C)}\right] \end{aligned} \tag{A.32}$$

At any point, the mass of movers in and out of A must equalize, so that $M_t^{W\rightarrow A} = M_t^{A\rightarrow W}$, implying:

$$\begin{aligned} (H_{t-1}^W - \phi_{t-1}^W) \cdot e^{-\lambda(\Omega_t + \delta \cdot ln(1+x_t^A))} + \phi_{t-1}^W \cdot e^{-\lambda\Omega_t} =& (\overline{H}^A - \phi_{t-1}^A) \cdot \left[1 - e^{-\lambda(\Omega_t - C + \delta \cdot ln(1+x_t^A))}\right] \\ &+ \phi_{t-1}^A \cdot \left[1 - e^{-\lambda(\Omega_t - C)}\right] \end{aligned} \tag{A.33}$$

Solving for the new occupancy cost, and replacing $\phi_t^A = \overline{H}^A x_t^A$, we obtain:

$$\boxed{\Omega^* = -\frac{1}{\lambda} \ln\left[\frac{\overline{H}^A}{e^{-\lambda\delta \cdot ln(1+x_t^A)} \cdot \left[1 + \overline{H}^A\,(1 - x_t^A) \cdot (e^{\lambda C} - 1)\right] + \phi + x_t^A \overline{H}^A \cdot (e^{\lambda C} - 1)}\right]} \tag{A.34}$$

Note that, at any point in time, occupation costs keep changing depending on the evolution of the local share of immigrants.

Ergodic Composition of Residents Let's again assume that all immigrants at $t = 0$ first arrive in W ($\phi_0^A = 0$). In the following periods, some of them they will start moving into A because of positive random draws of ε and an increasing availability of vacancies over the long run. The long-run dynamic equilibrium in ethnic shares must be such that the mass of *immigrants* entering and leaving A at any given period is equal:

$$\begin{aligned} M_t^{W\rightarrow A}(I) &= M_t^{A\rightarrow W}(I) \\ \phi_{t-1}^W \cdot e^{-\lambda\Omega_t} &= \phi_{t-1}^A \cdot \left[1 - e^{-\lambda(\Omega_t - C)}\right] \end{aligned} \tag{A.35}$$

Solving for Ω_t, we obtain:

$$\boxed{\Omega_t \;=\; -\frac{1}{\lambda}\ln\left[\frac{\phi^A_{t-1}}{\phi + \phi^A_{t-1}\cdot(e^{\lambda C}-1)}\right]} \tag{A.36}$$

Similarly, if the long-run ethnic composition of residents in A is stable, then it must also be true that the mass of *natives* entering and leaving A at any given period is the same:

$$\begin{gathered} M_t^{W\rightarrow A}(N) = M_t^{A\rightarrow W}(N) \\ (H^W_{t-1} - \phi^W_{t-1})\cdot e^{-\lambda(\Omega_t+\delta\cdot ln(1+x^A_t))} = (\overline{H}^A - \phi^A_{t-1})\cdot\left[1 - e^{-\lambda(\Omega_t - C + \delta\cdot ln(1+x^A_t))}\right] \end{gathered} \tag{A.37}$$

Solving for Ω_t, we now obtain:

$$\boxed{\Omega_t \;=\; -\frac{1}{\lambda}\ln\left[\frac{\overline{H}^A - \phi^A_{t-1}}{1+\left(\overline{H}^A - \phi^A_{t-1}\right)\cdot(e^{\lambda C}-1)}\right] - \delta\cdot ln(1+x^A_t)} \tag{A.38}$$

Setting Equation (A.36) equal to Equation (A.38), replacing $\phi^A = \overline{H}^A x^A$, and denoting by x^* the stable share of immigrants in A, we obtain:

$$x^*\left(1+(1-x^*)\overline{H}^A\left(e^{\lambda C}-1\right)\right) = (1-x^*)(1+x^*)^{\lambda\delta}\left(\phi + x^*\overline{H}^A\left(e^{\lambda C}-1\right)\right) \tag{A.39}$$

This equation implicitly defines the long-term immigrant share dynamic equilibrium. Note that the expression does not allow for an extreme scenario of full segregation, with either $x^A = 0$ or $x^A = 1$. This is a direct consequence of our setting, in which residents experience random draws of ε from an exponential distribution. Intuitively, even if residents have a strong aversion to immigrant presence, there is always a small probability that a resident will draw a substantially large ε, which may offset this aversion. However, the model can be parameterized to achieve nearly perfect segregation, with the proportion of natives converging to zero.

Note further that whenever $\delta = 0$, Equation (A.39) simplifies to: $x^* = \frac{\phi}{1+\phi}$, the proportion of immigrants in the overall population, thereby nesting **Result (A.5)**.

Result A.11 ***(Ergodic Immigrant Concentration under Distaste for Immigrants)*** *The ergodic share of immigrants in A, x^*, under native distaste for immigrants ($\delta > 0$), is always higher than in the scenario where natives are indifferent to them ($\delta = 0$).*

$$x^*_{\delta=0} < x^*_{\delta>0} \tag{A.40}$$

If $\delta > 0$, $\lambda > 0$, and $x^* > 0$, then $(1+x^*)^{\lambda\delta} > 1$. From equation (A.39), it follows that $(1+x^*)^{\lambda\delta} > 1$, which in turn implies:

$$x^* + x^*(1-x^*)\overline{H}^A \left(e^{\lambda C} - 1\right) > (1-x^*)\left(\phi + x^*\overline{H}^A \left(e^{\lambda C} - 1\right)\right) \tag{A.41}$$

Simplifying the expression above, we obtain:

$$x^*_{\delta>0} > \frac{\phi}{1+\phi} = x^*_{\delta=0} \tag{A.42}$$

To better visualize the relationship between native tastes for homophily and segregation, assume there are no moving costs, $C = 0$. Then, Equation (A.39) becomes,

$$x^* = \phi(1-x^*)(1+x^*)^{\lambda\delta} \tag{A.43}$$

For illustration purposes, normalize at $\lambda = 1$, and assume that $\phi = 0.11$. The latter number is the share of immigrants in Denmark. We can always find the x^* between 0 and 1 that numerically solves the expression above for different values of δ. Figure A.2 plots the ergodic immigrant share x^* as a function of the distaste parameter δ, with the horizontal dashed line indicating the benchmark immigrant share when natives have no ethnic preferences ($x^* = \frac{\phi}{1+\phi} = 11\%$). As the distaste parameter increases, the equilibrium immigrant share rises monotonically. With δs above 5, white flight intensifies: small increases in native aversion now generate disproportionately large amounts of immigrant segregation.

Scenario 5. Native Heterogeneity

Now, let's assume there are two types of natives: *Segregationists* and *Neutrals*. Native Segregationists (NS) dislike living with immigrants in A, as in Equation (A.23). Native Neutrals (NN) are indifferent to the presence of immigrants, and therefore have the same preferences as them, as given by Equation (A.2). Let's denote by θ the share of segregationists in the native population. Hence, if $\theta = 0$ we are back to scenario 3, whereas if $\theta = 1$ we are back to scenario 4. Also denote by θ^A_{NS} the measure of segregationists living in A, θ^A_{NN} the measure of neutrals living in A, θ^W_{NS} the measure of segregationists living in W, and θ^W_{NN} the measure of neutrals living in W.

Given our three groups, we can denote their respective mass of movers at period t as follows:

- Immigrants:

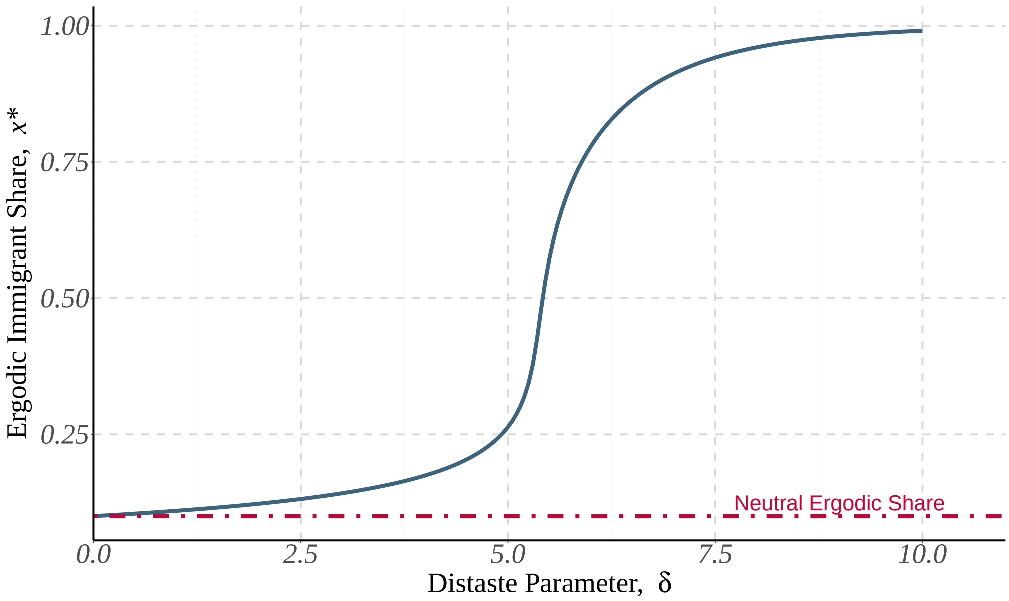


Figure A.2. Ergodic Immigrant Share and Distaste Parameter. This figure illustrates the ergodic immigrant share in A for different values of the distaste parameter, δ. The horizontal dashed line shows the benchmark share when natives are indifferent to immigration.

1. $M_t^{A\to W}(I) = \phi_{t-1}^A \cdot \left[1 - e^{-\lambda(\Omega_t - C)}\right]$
2. $M_t^{W\to A}(I) = \phi_{t-1}^W \cdot e^{-\lambda\Omega_t}$

- Native Segregationists:

1. $M_t^{A\to W}(NS) = \theta_{NS}^A \cdot \left[1 - e^{-\lambda(\Omega_t - C + \delta\cdot ln(1+x_t^A))}\right]$
2. $M_t^{W\to A}(NS) = \theta_{NS}^W \cdot e^{-\lambda(\Omega_t + \delta\cdot ln(1+x_t^A))}$

- Native Neutral:

1. $M_t^{A\to W}(NN) = \theta_{NN}^A \cdot \left[1 - e^{-\lambda(\Omega_t - C)}\right]$
2. $M_t^{W\to A}(NN) = \theta_{NN}^W \cdot e^{-\lambda\Omega_t}$

At each time period, the inflows and outflows must cancel each other out on average:

$$\begin{gathered} M_t^{A\to W} = M_t^{W\to A} \\ M_t^{A\to W}(I) + M_t^{A\to W}(NS) + M_t^{A\to W}(NN) = M_t^{W\to A}(I) + M_t^{W\to A}(NS) + M_t^{W\to A}(NN) \end{gathered} \tag{A.44}$$

Solving for Ω, we obtain:

$$\boxed{\Omega^* = -\frac{1}{\lambda}\ln\left[\frac{\overline{H}^A}{e^{-\lambda\delta\cdot ln(1+x_t^A)}\cdot\left[\theta_{NS}^W + \theta_{NS}^A e^{\lambda C}\right] + \left[\theta_{NN}^W + \theta_{NN}^A e^{\lambda C}\right] + \left[\phi^W + \phi^A e^{\lambda C}\right]}\right]} \tag{A.45}$$

It is straightforward to show that if there were no segregationists ($\theta_{NS}^{A} = \theta_{NS}^{W} = 0$), the equation above simplifies to Equation (A.19). Conversely, if all natives are segregationists ($\theta_{NN}^{A} = \theta_{NN}^{W} = 0$), it simplifies to Equation (A.34). In this scenario, the mix of neutral and segregationists will imply native different flows to and from A. Moreover, because $e^{-\lambda\delta \cdot ln(1+x_t^A)} < 1$, we can show that:

$$\Omega_t(x_t^A|\theta_{NN} = 0) < \Omega_t(x_t^A|\theta_{NS} \neq 0, \theta_{NN} \neq 0) < \Omega_t(x_t^A|\theta_{NS} = 0) \tag{A.46}$$

In words, for a given immigrant share in A, x^A, the larger the composition of segregationists in the economy, the smaller the equilibrium cost of living there, Ω.

The mix of neutral and segregationist natives allows for scenarios where the flight and avoidance behavior from segregationist natives can be partially (or even completely) offset by neutral natives. The ergodic equilibrium in ethnic composition can now be characterized by equalizing migration in and out of A across each of the three groups:

$$\begin{aligned} M^{A\to W}(I) &= M^{W\to A}(I), \\ M^{A\to W}(NN) &= M^{W\to A}(NN), \\ M^{A\to W}(NS) &= M^{W\to A}(NS). \end{aligned} \tag{A.47}$$

From these conditions, solving for Ω after some algebra we can obtain:

$$\begin{aligned} \Omega_t &= -\frac{1}{\lambda}\ln\Big[\frac{\phi^A}{\phi^W + \phi^A \cdot e^{\lambda C}}\Big] \\ \Omega_t &= -\frac{1}{\lambda}\ln\Big[\frac{\theta_{NN}^A}{\theta_{NN}^W + \theta_{NN}^A \cdot e^{\lambda C}}\Big] \\ \Omega_t &= -\frac{1}{\lambda}\ln\Big[\frac{\theta_{NS}^A}{\theta_{NS}^W + \theta_{NS}^A \cdot e^{\lambda C}}\Big] - \delta ln(1+x) \end{aligned} \tag{A.48}$$

For illustrative purposes, assume $C = 0$ again. Setting the three expressions equal to each other, we obtain:

$$\frac{\phi^A}{\phi^W + \phi^A} = \frac{\theta_{NN}^A}{\theta_{NN}^W + \theta_{NN}^A} = \frac{\theta_{NS}^A}{\theta_{NS}^W + \theta_{NS}^A} \cdot (1 + x^A)^{\lambda\delta} \tag{A.49}$$

Recall again that θ is the share of Segregationist natives, while $1 - \theta$ the share of neutral ones. Recall also that ϕ is the total measure of immigrants in the economy.

Therefore, we can rewrite the expression above as:

$$\frac{\phi^A}{\phi} = \frac{\overline{H}^A - \phi^A - \theta^A_{NS}}{1-\theta} = \frac{\theta^A_{NS}}{\theta} \cdot (1+x^A)^{\lambda\delta}. \tag{A.50}$$

Focusing on the last two terms, recall that $\phi^A = x^A\overline{H}^A$. Denoting the stable share of immigrants in A by x^*, and solving for θ^A_{NS}, we obtain:

$$\theta^A_{NS} = \frac{\theta\overline{H}^A(1-x^*)}{\theta + (1-\theta)(1+x^*)^{\lambda\delta}}. \tag{A.51}$$

Replacing θ^A_{NS} in Equation (A.50) with Equation (A.51), we find:

$$x^* = \frac{\phi(1-x^*)(1+x^*)^{\lambda\delta}}{\theta + (1-\theta)(1+x^*)^{\lambda\delta}}. \tag{A.52}$$

This equation implicitly defines the share of immigrants in equilibrium at A. Note that if $\theta = 1$ all natives are segregationists, and Equation (A.52) simplifies to Equation (A.43). On the other extreme, when $\theta = 0$, all natives are neutral and Equation (A.52) simplifies to $x = \frac{\phi}{1+\phi}$, the very proportion of immigrants in the overall population.

In addition, because $x^* > 0$, $\lambda > 0$, and $\delta > 0$, we must have that $(1+x^*)^{\lambda\delta} > 1$. This implies a positive association between the share of segregationists θ and the ergodic share of immigrants in A, x^*. *Ceteris Paribus*, the larger the number of neutral natives in the economy, the more they will offset the mobility responses of segregationists, and the smaller the segregation rates (i.e. the share of immigrants at A).

Now denote by $x^*_{\theta=0} = \frac{\phi}{1+\phi}$ the long-run share of immigrants in A under scenario 3 (with all-neutral natives), and denote by $x^*_{\theta=1} = \phi(1-x^*)(1+x^*)^{\lambda\delta}$ the same share under scenario 4 (with all-segregationist natives). Formally:

Result A.12 ***(Taste-Based Ergodic Sorting)*** *Under the scenario where natives are heterogeneous in their preferences towards immigrants, native flight is partially offset by native inflows. If $\delta > 0$, $\lambda > 0$, and $x^* > 0$, this implies that: $x^*_{\theta=0} < x^*_{0<\theta<1} < x^*_{\theta=1}$.*

$$\underbrace{\frac{\phi}{1+\phi}}_{x^*_{\theta=0}} < \underbrace{\frac{\phi(1-x^*)(1+x^*)^{\lambda\delta}}{\theta + (1-\theta)(1+x^*)^{\lambda\delta}}}_{x^*_{0<\theta<1}} < \underbrace{\phi(1-x^*)(1+x^*)^{\lambda\delta}}_{x^*_{\theta=1}} \tag{A.53}$$

Figures A.3 and A.4 illustrate how the mix of segregationists and neutrals shapes the local long-run immigrant shares. Figure A.3 plots the ergodic immigrant share x^* as a

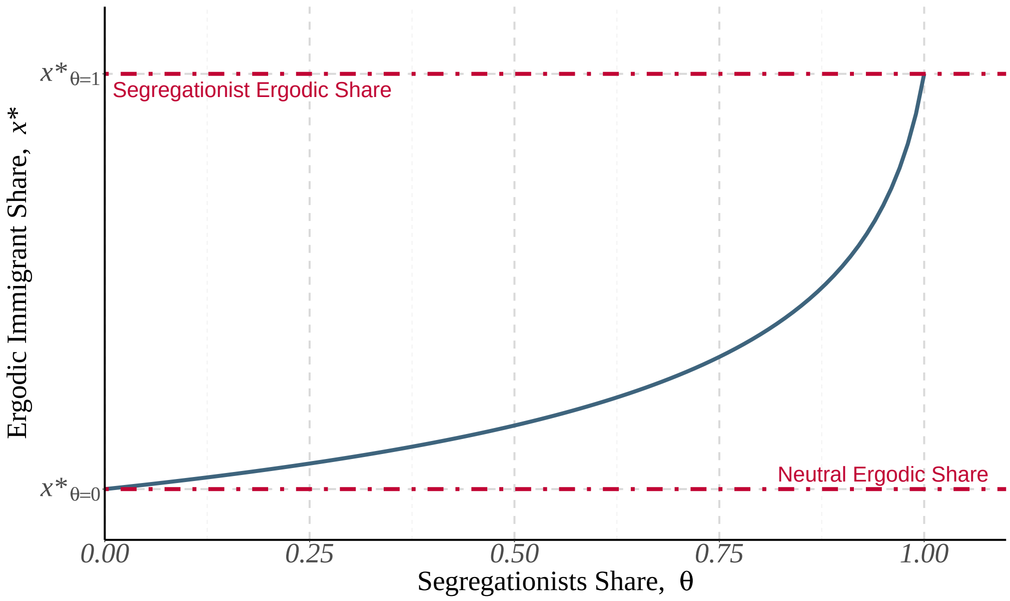


Figure A.3. Ergodic Immigrant Share and the Share of Segregationists. This figure illustrates the ergodic immigrant share in A for different values of the share of segregationist natives, θ.

function of θ, the fraction of segregationists in the native population. The relationship is monotonically increasing: as θ rises from 0 to 1, the equilibrium immigrant share increases from the neutral benchmark of $\frac{\phi}{1+\phi}$ to the segregationist benchmark, $\phi(1-x^*)(1+x^*)^{\lambda\delta}$. Figure A.4 extends this analysis by varying the intensity of segregationist preferences δ at three levels: low (light blue), medium (intermediate blue), and high (dark blue). The figure illustrates that the impact of the segregationist population share θ is amplified when the distaste is stronger.

Synthesis of Results and Implications

This appendix section has highlighted the pitfalls of relying on aggregate (often neighborhood-level) measures of population shares to study native behavior and residential mobility in response to immigration.

Micro-level panel data (i.e., individual- or household-level data) can be used to distinguish taste-based flight from replacement without displacement (i.e., high churn) or sorting by type among majorities. In short, while changes in neighborhood-level shares can be consistent with "flight" behavior by natives, these changes can also result from straightforward demographic churn or be obscured by compositional shifts within majority groups. Additionally, microdata has other advantages. For example, it has a better ability to identify household-specific shocks and heterogeneous treatment effects by type of immigrant or native. It can also distinguish the importance of interactions at different geographic levels and determine the destination of those who flee immigrant-dense

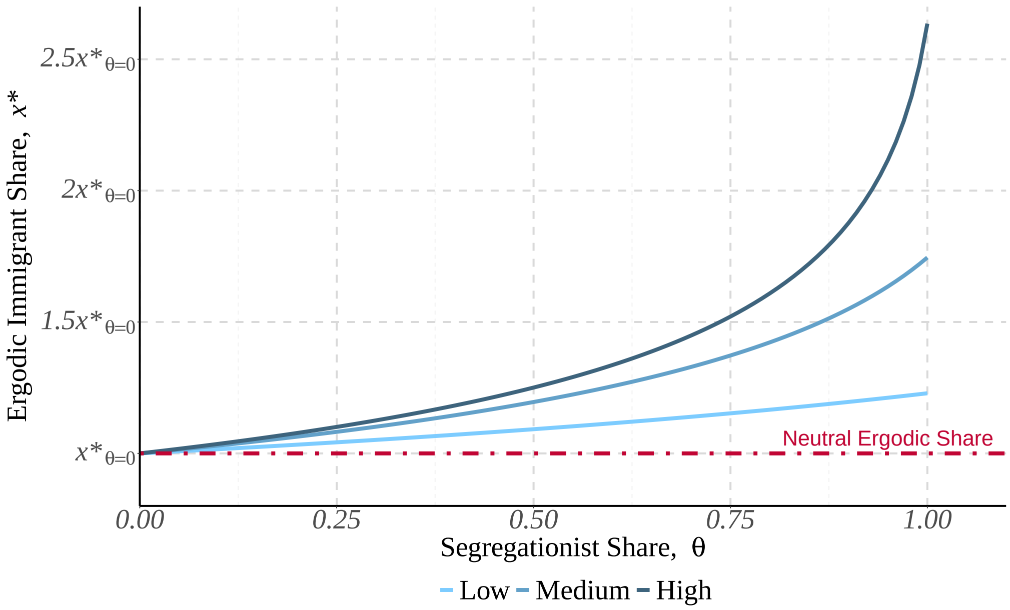


Figure A.4. Ergodic Immigrant Share and the Share of Segregationists, Across Levels of Distaste for Immigrants, δ. This figure illustrates the ergodic immigrant share in A for different shares of segregationists and across different levels for the distaste parameter, δ: low (light blue), medium (intermediate blue), and high (dark blue).

neighborhoods, as we do in the paper.

B Main Dataset and Additional Descriptive Stats

B1 Refugees in Denmark

Figure B.1 depicts the evolution of the composition of refugees in Denmark as a share in the total number of non-Western immigrants. We can see that in the 1980s, refugees were about 30% of immigrants, increasing to over 40% in the 1990s and remaining to that level until the early 2000s. After 2003, the share of refugees have been slowly decreasing, reaching around 35% of non-Western immigrants by the end of our period.

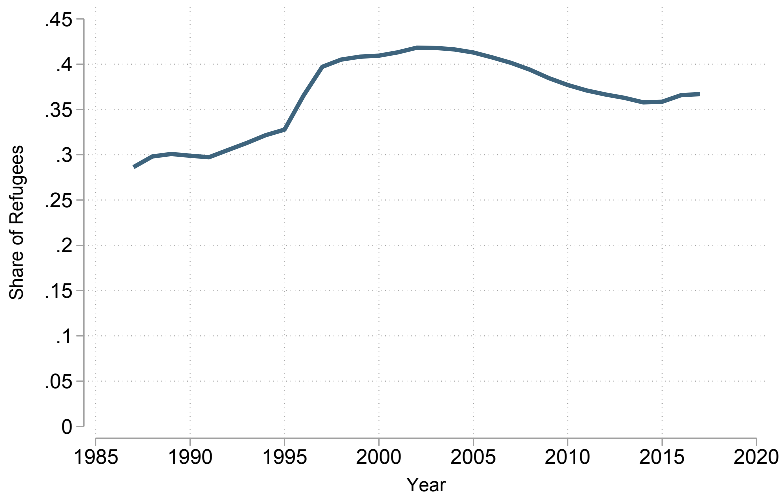


Figure B.1. Evolution of the Share of Refugees in Denmark. This figure illustrates the evolution of the share of refugees in Denmark in the total number of non-Western immigrants.

B2 Additional Maps

This section presents additional maps in order of appearance in the main text.

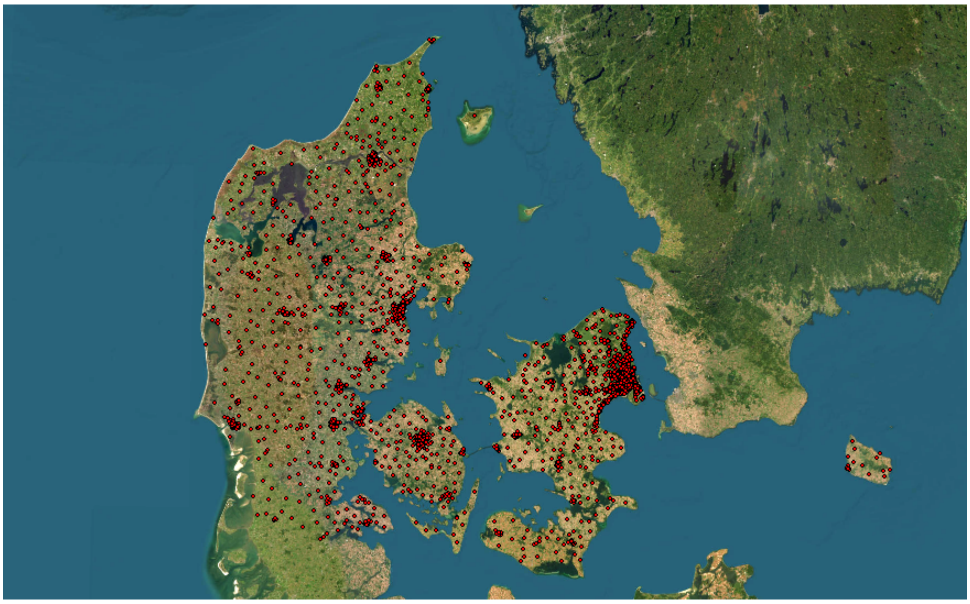


Figure B.2. Large Neighborhood Centroids. This figure illustrates the geographic distribution of neighborhoods across Denmark. The red points are the centroids of each area, which are based on the definitions of large neighborhoods as constructed by Damm and Schultz-Nielsen (2008).

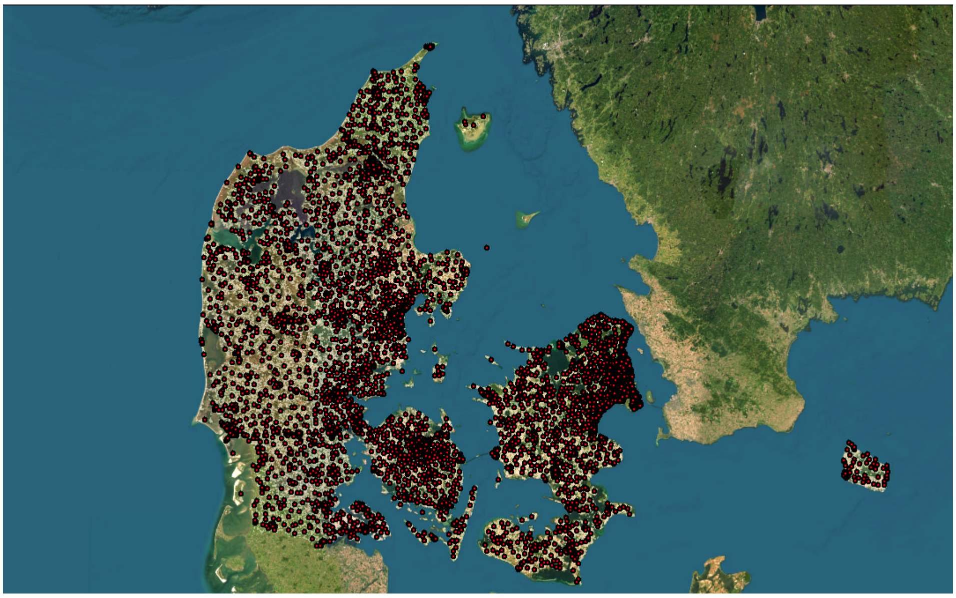


Figure B.3. Small Neighborhood Centroids. This figure illustrates the geographic distribution of small neighborhoods across Denmark. The red points are the centroids of each area, which are based on the definitions of small neighborhoods as constructed by Damm and Schultz-Nielsen (2008).

B3 Overlapping of Characteristics

In studying the heterogeneity in the response of natives due to immigrant presence, we find that there is a significant overlap in the number of heads of households that are living in public housing, multifamily buildings or complexes, and that are renters. Figure B.4 illustrates the relationship between the three variables by pooling the 30 years of available data. It shows that almost all residents (more than 99%) in public housing are also identified as renters. Moreover, the majority (99%) of the residents living in public housing are also inhabiting multifamily buildings. It shows that about 42% of the renters live in public housing and that 93% of renters live in multifamily buildings.

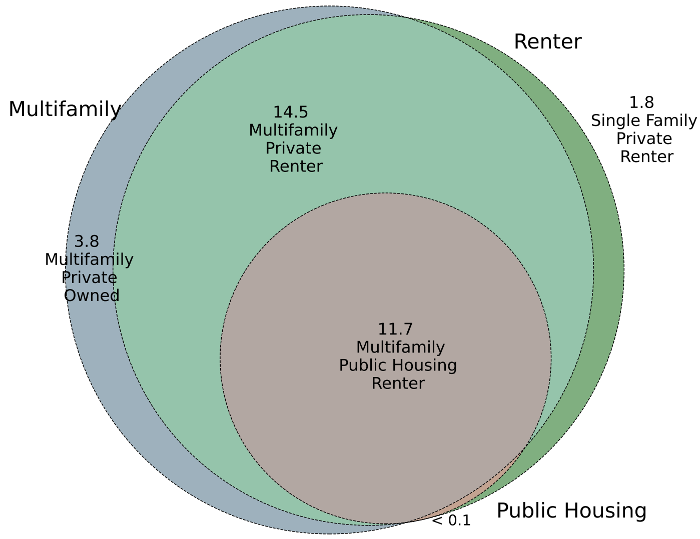


Figure B.4. Overlapping of Characteristics: Public Housing, Renters, and Multifamily Buildings. In this figure, we plot the Venn diagram to illustrate the relationship between living in multifamily buildings, in social housing, and being a renter. The numbers are in millions of heads of households in each group.

B4 Complete Distribution of Moving Distances

Figure 12 shows the density distribution of distances between the origin and destination of all Danish that moved to a different neighborhood between any two consecutive years. To facilitate the visualization, in the main text, we truncate the density plot to distances below 100 kilometers. Here, in Figure B.5, we present the complete distribution of distances. Again, we find a strong gravitational relationship that characterizes the moving distances.

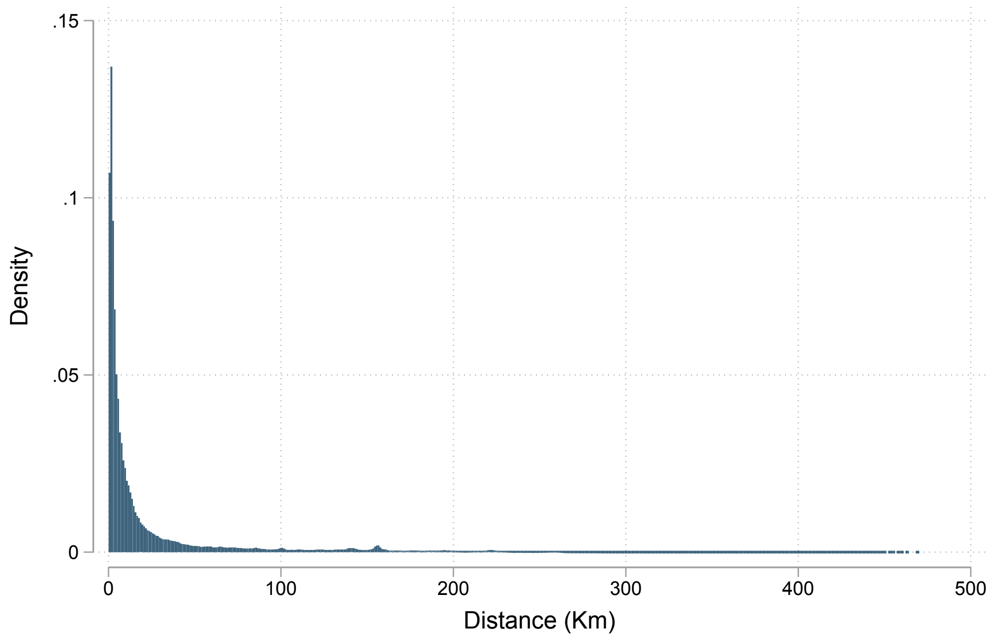


Figure B.5. Complete Histogram of Distances Between Origin and Destination Neighborhoods. In this figure, we plot the histogram of distances in kilometers between the origin and the destination neighborhoods for all Danish movers in our data.

B5 Definition of Refugees

All immigrants are classified as refugees or non-refugees depending on the year of their arrival to Denmark and their country of origin. The decision rules for the classification are listed below:

1. A country from where more than 50% of the persons arriving within a given year obtain asylum is classified as a refugee-country.

   For the sake of consistency over time, we further impose:

2. If the refugee-share exceeds 50% more than 10 years in the period 1997-2020, the country is classified as a refugee-country all years.

3. If the refugee-share exceeds 50% in year t and year $t + 2$, but not year $t + 1$, the country is classified as a refugee-country in year $t + 1$ as well.

4. If the refugee-share exceeds 50% in year $t + 2$, but not in: t, $t + 1$, $t + 3$ and $t + 4$, the country is not classified as a refugee-country in year $t + 2$.

Admission class information is available from 1997.[8] For the period prior to 1997 the classification is based on historical information about immigrants arriving to Denmark from conflicts and wars zones.

The following countries are classified as refugee-countries (selected arrival years): Afghanistan (all years), Angola (1999-2002), Armenia (1997-2008), Azerbaijan (all years), Bhutan (2008-2011), Bosnia and Herzegovina (1992-2004), Burundi (all years), Central African Republic (2014-2016), Colombia (2013-2014), DR of Congo (all years), Eritrea (all years), Ethiopia (1991-1998), Georgia (1998-2000), Indonesia (2003-2005), Iran (all years), Iraq (all years), Kosovo (1992-2011), Lebanon (1997-98), Libya (all years), Myanmar (all years), Poland(-1988), Republic of Yugoslavia (1992-2004, 2007-08), Rwanda (all years), Serbia (1992-2007), Serbia and Montenegro (1992-2007), Syria (all years), Somalia (all years), Sri Lanka (1983-2000), Stateless (all years), Sudan (all years), Vietnam (1975-1996), Yugoslavia (1992-2001).

[8] See, Statistics Denmark, Statbank, Table VAN8, https://www.statistikbanken.dk/van8.

C Urban Amenities

This appendix provides an overview of the construction of urban amenities for the years 1987 to 2017, as well as relevant descriptive statistics.

We follow Glaeser et al. (2001) and create three types of amenities. First, we construct several consumer amenities such as food services, entertainment, and groceries. To define consumer amenities, we utilize Danish administrative register data containing the entire universe of firms from 1987 to 2017. This data includes information about the exact location and the industrial sector code of each establishment, enabling us to link establishments to neighborhoods and calculate the concentration of amenities in each neighborhood for the past 30 years.[9] Amenities are determined by the total number of relevant establishments within a neighborhood. Figure C.1 illustrates the impressive 75% growth in the number of food services between 1987 and 2017. This encompasses restaurants, takeaways, pubs, bars, cafes, catering, and other food services. This steady highlights the importance of the food service industry, demonstrating how it has become an increasingly vital part of Danish society.

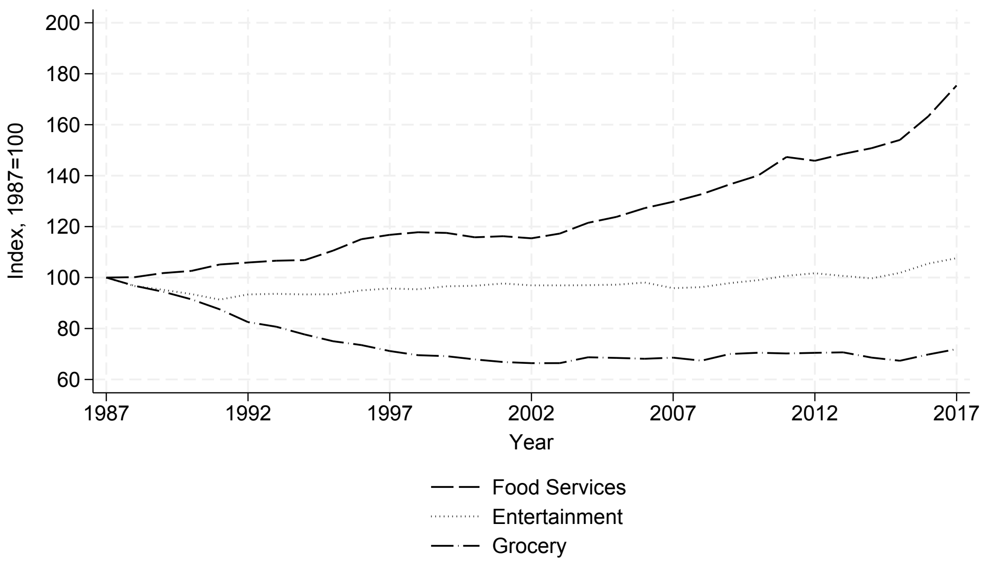


Figure C.1. Evolution of mean consumer amenities from 1987 to 2017. Consumer amenities are determined by the total number of relevant establishments within a neighborhood. The mean has been computed on a national basis.

[9]Between 1987 and 2017, Statistics Denmark has implemented multiple updates to their industrial sector codes. The first update occurred between 1991-1992 and resulted in a significant increase from 112 to approximately 800 codes. The second and third updates took place between 2003-2004 and 2007-2008, respectively. To ensure consistency throughout the years, we manually compared the industrial codes prior to 1992 with those after 1991 to find a corresponding match. We repeated this process for each period of modification implemented by Statistics Denmark.

We also construct two additional consumer amenities: entertainment and grocery stores. Entertainment is defined broadly and includes over 30 industrial sectors codes such as activities of sports clubs, amusements, and recreation activities, theaters and concerts, news agency operations, as well as motion picture and video production.[10] Figure C.1 shows that the entertainment amenities have slightly decreased during the 90s compared to the reference year 1987, but since the end of the 90s, there has been a small but more or less constant increase. Consequently, the number of entertainment amenities has increased by 7% in 2017 compared to 1987. Grocery stores including supermarkets, discount stores, and retail sales of groceries and late-night stores have decreased by 30% compared to 1987. This is related to the fact that many small stores disappeared while large retailers and shopping malls emerged. Table C.1 shows summary statistics for the three consumer amenities on the neighborhood level, pooling all years together. There is a large variation in each amenity across neighborhoods, nevertheless, the average presence of the amenities in the neighborhood is around one.

Table C.1. Summary statistics for consumer amenities

	Mean	SD	Min	Max
Food services	1.229	14.035	0	1,712
Entertainment	0.739	6.205	0	750
Grocery	0.529	6.789	0	1,091

Note: Number of observations is 261,272 (years × neighborhoods). The unit of measurement is number of establishments.

The second set of amenities refers to public goods such as primary schools, doctors, and banks. We use the same register data to define these amenities, and primary schools include pre-primary, primary, and secondary levels, special schools for disabled persons, and youth and continuation schools.[11] Figure C.2 shows that the average number of primary schools in a neighbourhood has decreased by 15% over the last three decades compared to 1987. Doctors provide 13 different healthcare services, including general medical and specialist practice activities, psychological guidance, dental practice, and chiropractors - and their average number in a neighbourhood has increased by 34% in Denmark since 1987. In contrast, the average number of banks in a neighbourhood has decreased by almost 70%, as revealed in Figure C.2. This is due to the centralization of the banking sector. Table C.2 presents the summary statistics for amenities representing public goods, pooled across all years. As with Table C.1, there is significant variation in

[10] You can find the full list of industrial codes defining entertainment amenities prior to 1992 and between 1992-2006 in Statistics Denmark (1995), and from 2006 onwards in Statistics Denmark (2007).

[11] In Denmark, it is common for children to attend youth and continuation schools for one or two years after primary school to gain more professional skills before moving on to secondary school.

each amenity across neighborhoods. On average, there is less than one bank and primary school per neighborhood, but about one doctor per neighborhood.

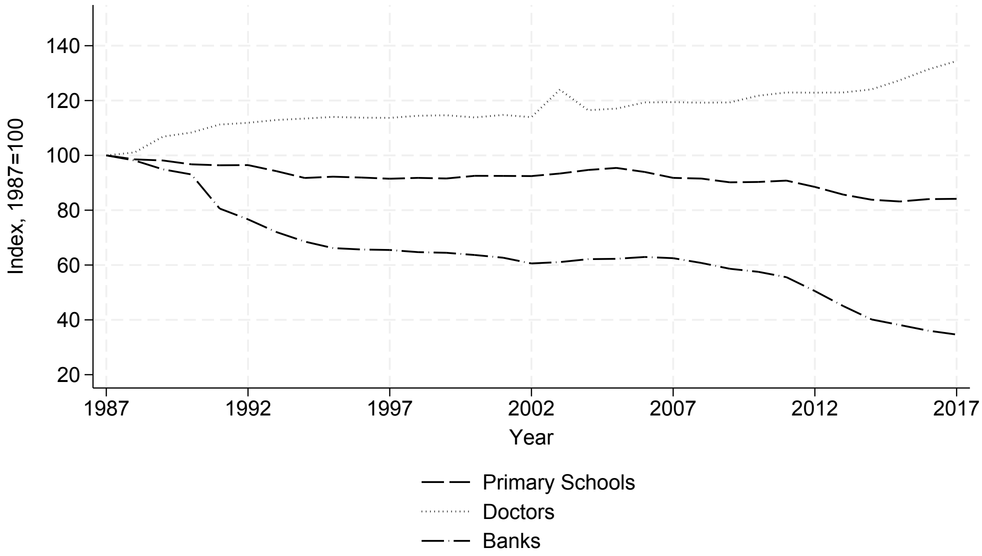


Figure C.2. Evolution of mean of selected public goods from 1987 to 2017. Consumer amenities are determined by the total number of relevant establishments within a neighborhood. The mean has been computed on a national basis.

Table C.2. Summary statistics for public goods

	Mean	SD	Min	Max
Primary school	0.348	3.482	0	552
Doctors	0.910	8.013	0	895
Banks	0.262	3.072	0	562

Note: Number of observations is 261,272 (years × neighborhoods). The unit of measurement is number of establishments.

Third, we calculate the proximity of each neighborhood to various transportation facilities, such as motorways, railway train stations, airports, and high voltage lines, as well as to various natural amenities, such as coasts, forests, lakes, recreational areas, statues, and cemeteries.[12] To do this, we measure the distances in meters from the center of each neighborhood to the relevant facilities and amenities.

Geographical maps and the evolution of the discussed amenities can be viewed and explored at http://www.ditnabolag.dk.

[12]In Denmark, cemeteries are often utilized as public parks.

D Additional Empirical Results

This section presents additional empirical results in order of appearance in the main text.

D1 Robustness to the Neighborhood Definition

In this section, we assess the robustness of the estimates from the Table 3 to the definition of neighborhood. As described in Section 4, we adopt the definition of the larger neighborhood from Damm and Schultz-Nielsen (2008) in our main analysis. In this section, we present the equivalent estimates of the baseline results using the small neighborhood definition from Damm and Schultz-Nielsen (2008). Table D.1 presents the estimates, showing that our findings are robust to the choice of definition for neighborhoods. One difference is the coefficients for the share of immigrants in the building complex in column (4) become statistically non-significant mainly because the fixed effects for neighborhood-year and building complex soak up most of the variation since there is a large overlap between building complexes and the smaller neighborhood definitions.

Table D.1. Regressions using the smaller neighborhood definitions

	Dependent Variable: Moved out of the Neighborhood in $t+1$			
	Full Sample		Large Building or Complex (10 or more units)	
	(1)	(2)	(3)	(4)
Immigrant share (neighborhood)	0.043***	0.024***	0.036***	
	(0.005)	(0.004)	(0.005)	
Immigrant share (building complex)				0.005
				(0.004)
N	53,332,175	53,332,175	14,424,608	14,423,944
R^2	0.033	0.097	0.089	0.110
Controls		✓	✓	✓
Region × Year FEs	✓	✓	✓	
Neighborhood FEs	✓	✓	✓	
Neighborhood × Year FEs				✓
Building Complex FEs				✓

Notes: The observations consist of heads of households, non-immigrants, with ages above 18 years old. Standard errors in parentheses are clustered by neighborhood. Statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

D2 Robustness to the Definition of Building Complex

In Section 5, when focusing on the building complexes, we determined that building complexes with fewer than 10 apartment units should be excluded from the sample, to avoid issues with excess volatility in the share of immigrants. One natural concern is whether the results are robust to alternative choices for the minimum number of apartment units within a building complex. In this section, we estimate the equivalent specifications as columns (4) and (5) from Table 3 but determine alternative cut-offs for the minimum number of apartment units in a building complex. Specifically, we estimate for building complexes with more than 5 apartments (columns 1 and 2), 10 apartments (columns 3 and 4), 15 apartments (columns 5 and 6), and 20 apartments (columns 7 and 8). Results are in Table D.2 and show that the coefficients are robust to the choice of the minimum number of apartments.

Table D.2. Robustness to the Number of Units Cutoff

	Dependent Variable: Moved out of the Neighborhood in $t+1$							
	More than 5 units		More than 10 units		More than 15 units		More than 20 units	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Immigrant share (neighborhood)	0.039***		0.036***		0.034***		0.031***	
	(0.007)		(0.008)		(0.008)		(0.008)	
Immigrant share (building complex)		0.011***		0.015***		0.019***		0.023***
		(0.003)		(0.004)		(0.004)		(0.005)
N	16,499,478	16,498,847	14,612,584	14,612,240	13,480,753	13,480,548	12,260,093	12,259,946
R^2	0.088	0.107	0.084	0.101	0.082	0.098	0.082	0.097
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Region × Year FEs	✓		✓		✓		✓	
Neighborhood FEs	✓		✓		✓		✓	
Neighborhood × Year FEs		✓		✓		✓		✓
Building Complex FEs		✓		✓		✓		✓

Notes: The observations consist of heads of households, non-immigrants, with ages above 18 years old. Standard errors in parentheses are clustered by neighborhood. Statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

D3 Robustness to the Inclusion of Trends Associated with Shares

In this section, we test the sensitivity of our OLS estimates to the inclusion of two terms. We basically estimate Equation (11), including two additional control variables. One is the share of refugees living in the neighborhood in 1987 times a five-year trend variable, and the second is the interaction between the share of public housing located in the neighborhood time the five-year trend variable. The idea is that these interaction terms capture the long-term trends in moving rates associated with the 1987 distribution of public housing and refugees across neighborhoods in Denmark. Table D.3 presents the estimates equivalent to specifications (2)-(5) from Table 3. The estimates are very similar to the ones in Table 3, suggesting that our OLS results are robust to controlling for the trends associated with the exposure measures we use to build the ASM-IV.

Table D.3. Robustness to the inclusion of additional trend variables

	Dependent Variable: Moved out of the Neighborhood in $t+1$			
	Full Sample	Single Family	Small Building or Complex (less than 10 units)	Large Building or Complex (10 or more units)
	(1)	(2)	(3)	(4)
Immigrant share (neighborhood)	0.017***	−0.001	0.003	0.032***
	(0.006)	(0.006)	(0.011)	(0.008)
N	53,332,175	27,978,943	10,740,642	14,612,584
R^2	0.092	0.050	0.108	0.084
Controls	✓	✓	✓	✓
Region × Year FEs	✓	✓	✓	✓
Neighborhood FEs	✓	✓	✓	✓

Notes: The observations consist of heads of households, non-immigrants, with ages above 18 years old. Standard errors in parentheses are clustered by neighborhood. Statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

D4 Robustness to the Exclusion of "Ghetto" Neighborhoods

Starting in 2010, Denmark implemented a large-scale intervention on neighborhoods with high concentrations of immigrants, named the "Ghetto Programme". The policy consisted of identifying disadvantaged neighborhoods and implementing a series of policies to alter the trajectory of these areas, especially regarding the concentration of immigrants (Damm et al., 2025). Among other criteria, neighborhoods with more than 50% immigrant residents were included on the "Ghetto List", and were targeted with a variety of social programs and investments in local infrastructure. Damm et al. (2025) shows that the policy had effective results in reducing crime in the treated areas.

The intervention involved multiple goals, including changing the demographic profile of the areas. Therefore, in this section, we test the sensitivity of our estimates to the exclusion of neighborhoods potentially treated in the "Ghetto Programme". The main concern is that our estimates are sensitive to the program. We run the same specification as in Table 3, but excluding neighborhoods with high shares of non-Western immigrants after 2010. The estimates presented in Table D.4 show that our estimates do not change in response to the exclusion of "Ghetto" areas, suggesting that our results are not

Table D.4. Robustness to the Exclusion of "Ghetto" Neighborhoods

	Dependent Variable: Moved out of the Neighborhood in $t+1$					
	Full Sample		Single Family	Small Building or Complex (less than 10 units)	Large Building or Complex (10 or more units)	
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant share (neighborhood)	0.040***	0.022***	–0.002	0.004	0.036***	
	(0.009)	(0.006)	(0.005)	(0.010)	(0.008)	
Immigrant share (building complex)						0.015***
						(0.004)
N	53,289,741	53,289,741	27,977,693	10,737,111	14,574,931	14,574,587
R^2	0.028	0.092	0.050	0.108	0.084	0.101
Controls		✓	✓	✓	✓	✓
Region × Year FEs	✓	✓	✓	✓	✓	
Neighborhood FEs	✓	✓	✓	✓	✓	
Neighborhood × Year FEs						✓
Building Complex FEs						✓

Notes: The observations consist of heads of households, non-immigrants, with ages above 18 years old. Standard errors in parentheses are clustered by neighborhood. Statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

D5 Robustness to the Exclusion of Own Building Exposure

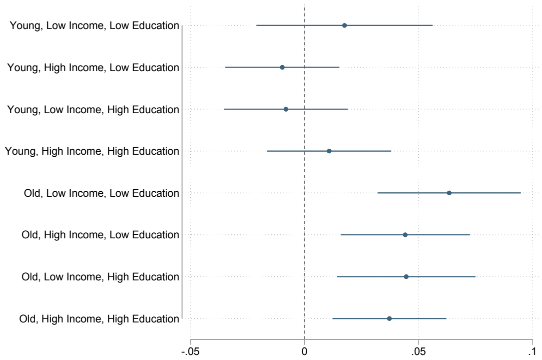
While our baseline specifications include a comprehensive set of controls and fixed effects for individual, household, building, and neighborhood characteristics, a potential concern remains regarding mechanical effects that could bias our estimates. Specifically, unobserved building-specific factors may simultaneously influence both immigrant settlement patterns and native mobility decisions. In this section, we test the robustness of our results by excluding immigrant exposure at the individual's own building level from our key explanatory variables. Specifically, we recalculate the immigrant shares at both the neighborhood and building complex levels after removing all individuals residing in the same building as each household head. This adjustment ensures that our estimated effects capture responses to immigrant presence in the broader local area rather than mechanical correlations arising from own-building composition.

Table D.5 presents the estimates using these adjusted immigrant shares. The results remain largely consistent with our baseline findings in Table 3, confirming that native flight is not driven by mechanical effects from own-building exposure. The coefficients are slightly attenuated but remain statistically significant and economically meaningful.

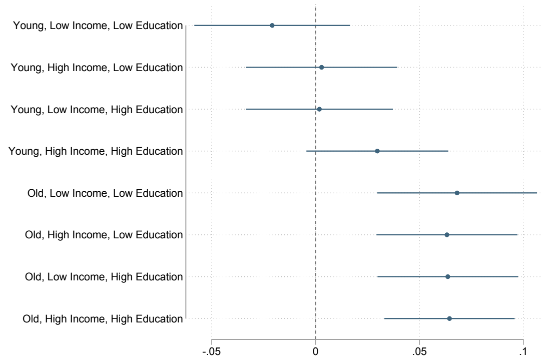
Table D.5. Robustness to the Exclusion of Own Building Exposure

	Dependent Variable: Moved out of the Neighborhood in $t+1$				
	Full Sample		Single Family	Small Building or Complex (less than 10 units)	Large Building or Complex (10 or more units)
	(1)	(2)	(3)	(4)	(5)
Immigrant share (neighborhood)	0.031***	0.017***	–0.003	0.002	0.029***
	(0.009)	(0.006)	(0.005)	(0.010)	(0.008)
N	53,332,170	53,332,170	27,978,938	10,740,642	14,612,584
R^2	0.028	0.092	0.050	0.108	0.084
Controls		✓	✓	✓	✓
Region × Year FEs	✓	✓	✓	✓	✓
Neighborhood FEs	✓	✓	✓	✓	✓

Notes: The observations consist of heads of households, non-immigrants, with ages above 18 years old. Standard errors in parentheses are clustered by neighborhood. Statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.



(a) Sample set 1



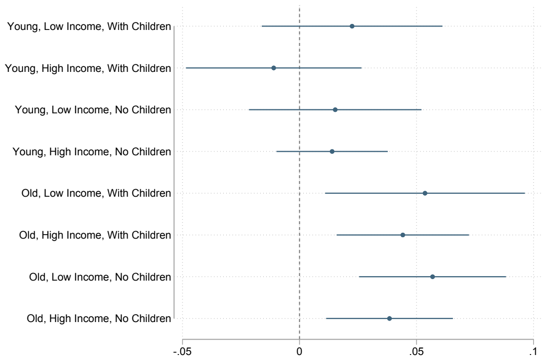
(b) Sample set 2

Figure D.1. Heterogeneity Analysis by Age, Education, and Income Splitting the Sample Randomly into Two Parts. This figure presents the heterogeneity analysis as discussed in Section 5.6.1, but randomly splitting the sample into two parts.

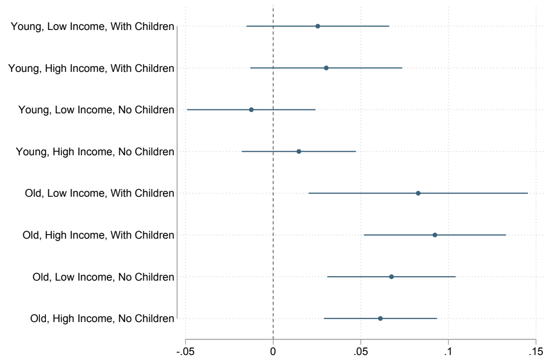
D6 Robustness of heterogeneity results

When conducting our heterogeneity analysis in Section 5.6, we split the sample randomly into two distinct subsets. This randomization is executed at the neighborhood level. The selection of neighborhoods into each group is determined by their alternating ranks, which are established by sorting them based on their unique identification numbers. The same regression analysis is then independently applied to each subset.

Figures D.1, D.2 and D.3 plot the estimated coefficients for the share of immigrants and the interaction terms with each dummy variable that categorize the groups as described in Section 5.6, but for two random partitions of the sample.

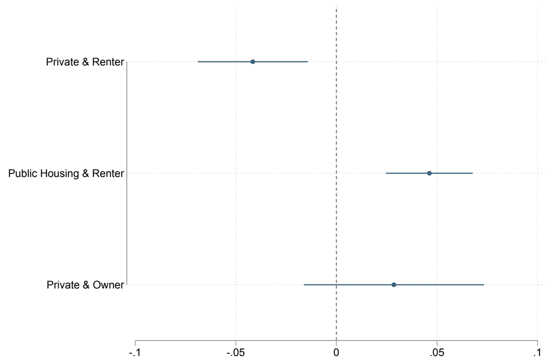


(a) Sample set 1

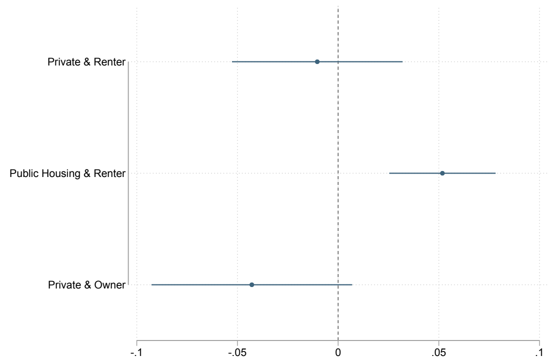


(b) Sample set 2

Figure D.2. Heterogeneity Analysis by Age, Income, and Children in the Household Splitting the Sample Randomly into Two parts. This figure presents the heterogeneity analysis as discussed in Section 5.6.2, but randomly splitting the sample into two parts.



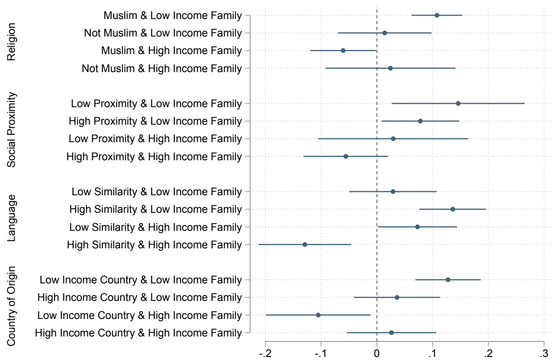
(a) Sample set 1



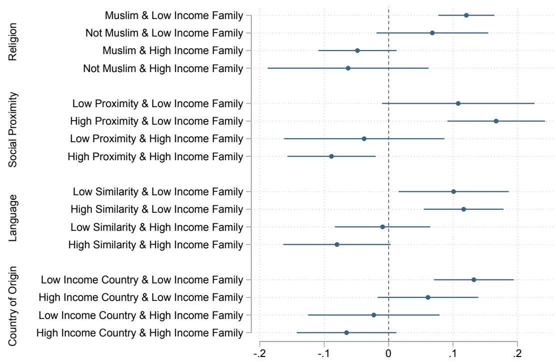
(b) Sample set 2

Figure D.3. Heterogeneity Analysis by Homeownership Status Splitting the Sample Randomly into Two parts. This figure presents the heterogeneity analysis as discussed in Section 5.6.3, but randomly splitting the sample into two parts.

Figure D.4 plot the estimated coefficients for heterogeneity analysis to non-Western immigrant sub-groups.



(a) Sample set 1

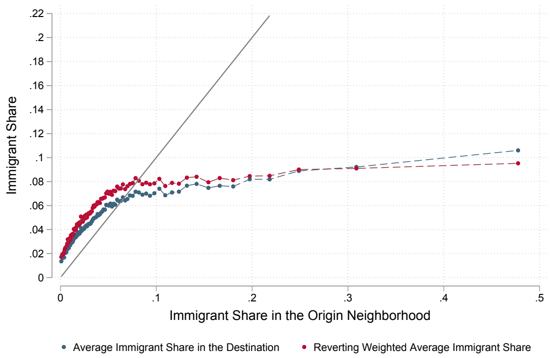


(b) Sample set 2

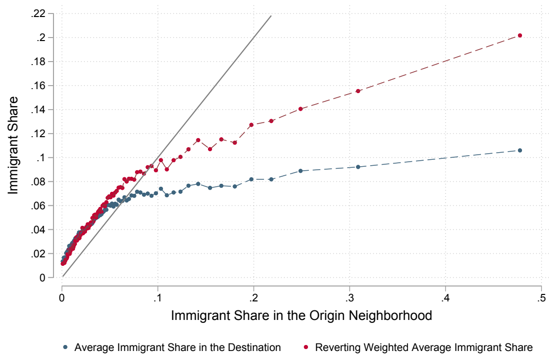
Figure D.4. Heterogeneity Analysis to Non-Western Immigrant Sub-groups by Splitting the Sample Randomly into Two Parts. This figure presents the heterogeneity analysis as discussed in Section 6, but randomly splitting the sample into two parts.

D7 Reverting Weighed Average Immigrant Share Under Different Spatial Decay Parameter Values

From Section 8, we propose the construction of the Reverting Weighed Average Immigrant Share as a good measure to serve as a comparison to actual patterns of destination choices for natives moving to a different neighborhood. In the main text, we adopt a spatial decay parameter of 2 to calculate the reverting weighted average immigrant share. In this section, we explore the patterns when different values for the spatial decay parameter are used. In Figure D.5, we plot a similar figure as in Figure 12 with the nonparametric binned scatter plots of average immigrant shares at the destination (blue) and reverting weighed average immigrant share (red) versus the mean immigrant share at the origin. In panel (a) we use $\beta = 1$, while in panel (b) we use $\beta = 3$. We find a similar pattern from Figure 12, where the actual averages (blue) are below the expected averages (red), suggesting that the movers are choosing to move to destinations with lower shares of immigrants relative to the baseline mean reversion scenario. An important takeaway from this section is that the faster we allow for the weights to go to zero as the distance increases, the larger the gap between the two curves.



(a) $\beta = 1$



(b) $\beta = 3$

Figure D.5. Reverting Weighed Average Immigrant Share with Different Coefficients for the Spatial Decay Parameter, β.