Immigrants and Native Flight: Geographic Extent and Heterogeneous Preferences*

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

June 28, 2024

Abstract

Is ethnic segregation in Europe driven by native flight or immigrant self-isolation? If the former, which natives avoid immigrants? Which immigrants? What is the geographic scope of homophilic residential preferences? We answer these questions using a matched panel containing the universe of individuals and properties in Denmark from 1987 through 2017. We take advantage of the quasi-random nature of refugee placements and simulated exogenous Markov-chain predictions to generate experimental variation regarding local immigrant arrivals. We find strong evidence of native flight, even at the building level. Flight is stronger among the old and a reaction to the arrival of low-income immigrants. As neighborhoods become more immigrant-dense, housing prices decline, and subsequent move-ins are more likely to be other immigrants or young, low-income native citizens without children.

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

*The authors thank Statistics Denmark for providing the data. We would like to thank conference and seminar participants at ROCKWOOL Foundation Research, the University of South Florida, Universitat Autònoma de Barcelona, Boise State University, the University of Bern, the Urban Economics Association North American and European Meetings, the 40th ARES Annual Conference, and the 13th Annual Conference on "Immigration in OECD Countries" 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.
¶MIT, Center for Real Estate, Department of Urban Studies and Planning, USA, email: vsantann@mit.edu.
‖ROCKWOOL Foundation Research, DK, email: mls@rff.dk.

1 Introduction

Over the past decades, Western European countries have experienced an uptick in immigration. According to the UNHCR, approximately 1 million immigrants reach the region annually, generating heated debates about their accommodation, assimilation, and impacts on destination economies. In fact, immigration has become *the* most salient political issue in the EU and other developed countries. Anti-immigrant feelings are also rising in developing countries—such as Turkey, Tunisia, South Africa, and Colombia.

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. Conversely, 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) argues 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 ethnic group in a neighborhood and the educational attainment as well as earnings of the second-generation ethnic group. The study suggests that as the skill level of immigrants increases, their children's educational outcomes and income levels consequently improve.

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) shows that native exits from immigrant-dense metro areas are not random in France: 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 extended past research onto neighborhood outcomes rather than simply comparing 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 native residents (Miyao, 1979; Bond and Coulson, 1989; Saiz and Wachter, 2011; 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 decline in the growth of the white population and complete segregation (Schelling, 1971; Bond and Coulson, 1989; Frankel and Pauzner, 2002; Card et al., 2008). On the other hand, racial segregation may also be partially attributed to the perception of lower-quality public services and urban amenities in such areas (Bayer et al., 2007).[3]

Research in Economics about ethnic preferences and the drivers of neighborhood segregation in European cities 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

[2]Papers in this literature are too numerous to itemize, 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); Derenoncourt (2022)

[3]Saiz and Wachter (2011); 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, e.g., Gutierrez-i-Puigarnau et al. (2016) and Mulalic and Rouwendal (2020).

flows to and from urban neighborhoods, Accetturo et al. (2014), Moraga et al. (2019) and Andersson et al. (2021) show that immigrant inflows tend to be associated with net native outflows in Italy, Spain, and Sweden. 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]

In this paper, we use comprehensive administrative data for the entire population of Denmark, providing micro-geographic information for over three decades—which annually identifies 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 in Europe 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 that help us identify clear heterogeneous treatment effects and that are ultimately important for policy in a European context.

First, we explicitly model the micro-individual probability of natives moving out of a neighborhood as a function of the local share of immigrants. *Conceptually*, modeling individual probabilities allows us for a true revealed-preference test of native preferences. Tipping-point and outflow-to-inflow estimates are typically based on changes in net ethnic shares by neighborhood. If the native flight of prejudiced natives is countered by commensurate arrivals of tolerant ones, studies based on net flows will miss the existence of discriminatory preferences.[5] Our approach also avoids reverse causation from native exits to subsequent immigrant arrivals via increased vacancies. This problem is difficult to avoid using decennial aggregate population changes.

Empirically, modeling individual probabilities allows us to control for a very rich set of household characteristics and local amenities that remained unobserved in previous studies—focusing on neighborhood dynamics—and which may confound the relationship of interest. For instance, immigrant-dense neighborhoods may also be hosting natives with high propensities for mobility in general. In turn, highly available vacancies may make it easier for immigrants to move in.[6]

[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.

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

[6] This is a similar problem than encountered in the study of gentrification because low-income households have large churn propensities to start with, so it is unclear if they were truly displaced. Pennington

Second, we employ alternative identification strategies that are highly consistent with causal interpretations. Endogeneity challenges plague the estimation of outflow-to-inflow parameters (Moraga et al., 2019). On the one hand, contemporaneous shocks to a neighborhood may account for both the association between changes to its immigrant density and the propensity of natives to move out. On the other hand, the types of neighborhoods where immigrants clustered in the past may have experienced differential long-term trends, even in the absence of further inflows. Standard IV strategies using shift-share predictions of local immigrant arrivals address the former concern but may exacerbate the latter. We, therefore, provide estimates based on a novel Arrival-Stayer Markov Instrumental Variable (ASM-IV). Our ASM-IV combines random assignment of newly arriving refugees with mechanical mobility estimates. We also devote effort to minimize concerns of over-reliance on a few initial shares raised by Bartik instruments (Goldsmith-Pinkham et al., 2020), by separating the variation arising from random arrivals. In addition, our estimates control for neighborhood fixed effects, an exhaustive set of time-changing individual and neighborhood variables, and for trends that differ according to the initial characteristics of the area.

Third, we consider different housing typologies, as previous research has documented differences in overall segregation across urbanistic environments (Salazar, 2020). We document significant differences in behavior: native flight is stronger between multifamily buildings sited in denser areas. We estimate precise zero effects within the subsample of single-family homes. While the distance between households in suburban typologies keeps people apart (Putnam, 2000), it may also mitigate potential social tensions. Conventional local average treatment effects (LATE) obtained via shift-share IVs may overestimate white/native flight. This is so because the neighborhoods that hosted minorities in the past tend to be in central dense areas, precisely where white/native out-migration response is the 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., 2022). However, in Northern and Eastern Europe—as in many other parts of the world—the relevant micro-unit for social interaction is the housing complex: a collection of adjacent 3-4 floor-tall buildings with shared spaces, which is perceived as the relevant living environment for residents.[7] In addition to providing

(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."*

[7] Similar mid-rise, horizontally-expansive housing typologies are prevalent in some parts of the USA,

evidence on micro-interactions, these specifications allow us to 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 an analysis of nonlinearities (tipping points). Conventional tipping point estimates focus on *net* changes in the overall share of majority populations and are therefore unable to identify churn among the white/native population.

Sixth, because we focus on revealed preferences, we can identify the marginal natives who changed their behavior. We observe strong evidence of heterogeneity: households led by older individuals tend to exhibit more aversion 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 flee relative to homeowners or renters in private units.

Seventh, we 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.

Eighth, we show that—in areas with high immigrant shares—subsequent incoming residents are more likely to be other non-Western immigrants or low-income Danish citizens who are young and have no children. Overall, our results show that—in addition to monitoring local ethnic shares—policymakers may also need to worry about sorting and changes in the local composition of the native population in minority areas.

Ninth, previous research has not been able to characterize the destinations of leavers. We confirm that—as expected under the hypothesis of ethnic preferences—natives who moved out of areas with high immigrant concentrations chose destinations with relatively smaller immigrant shares.

Tenth, to dispel concerns about publication bias—thanks to the high power offered by our dataset—we show the results to be robust to a split sample strategy, whereby we randomly allocate neighborhoods into two alternate sub-samples. Split sample strategies also confirm the credibility of our analyses of heterogeneous treatment effects, because we could see some effects arise at random as one estimates parameters for an increasing number of subpopulations.

Finally, we combine all of the earlier findings with evidence on housing prices. Lower housing prices are sufficient, but not necessary, to demonstrate the existence of ethnic preferences. However, they are neither necessary nor sufficient for the decline of the local

such as courtyard apartments in Chicago (Bluestone, 2017) or garden apartments in California and the Southwest (Hise, 1995).

native population. By weaving together findings on prices and household mobility, we provide a complete picture of the process of ethnic segregation in a European context.

The remainder of the paper is organized as follows. In Section 2, we explain the data and provide descriptive statistics. Our estimation strategy is introduced in Section 3. Section 4 presents the main empirical results while Section 5 discusses heterogeneous responses to different immigrant groups. In Section 6, we analyze how the share of non-Westerners influences the characteristics of new residents. Section 7 assesses the destination choices of people who moved out, while Section 8 studies housing price effects. Section 9 concludes.

2 Data

2.1 Statistics Denmark: a Population Registry

We take advantage of 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 were allowed to link a number of individual demographic variables—such as 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.[8] 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.

2.2 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

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

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.[9] Nonetheless, Table C.1 in Appendix C 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).

2.3 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. Descendants are people born in Denmark but with both immigrant parents. 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. Because of 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 *entry door* by building. We categorize single-family detached buildings as structures with a unique *entry door*—which coincides with

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

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 such complexes: collections of identical, 3-6 story-high buildings attached horizontally. In the figure—extracted from Google Maps©—we appreciate 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 Metro 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 neighborhood*. Conservatively, moves within a neighborhood are not classified as native flight. Nevertheless, we compute three alternative dummies for moving out of an address, building, and neighborhood (shown in the first three rows of Table 1).

11% o 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% of them 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 tended to stay 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. A third of households live in large building com-

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	0.08	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	0.42	0
Higher education	0.244	0.429	0
Vocational education	0.37	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	0.59	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	2.82	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 (meters)	4,973.174	5,756.858	2,449.693
Proximity to roads (meters)	15,399.663	17,690.734	11,912.748
Proximity to the coast (meters)	6,368.713	8,252.439	3,191.5
Proximity to waste (meters)	26,949.011	23,230.448	17,880.830
Proximity to lake (meters)	409.863	317.420	343.853
Proximity to forest (meters)	325.86	335.282	238.03
Proximity to train station (meters)	4,526.765	12,796.138	1,417.171
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).

plexes (with more than ten units), and approximately 19% 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).[10]

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 at 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 0.3%, again with a high standard deviation. This implies that non-Western immigrants disproportionately live in multi-unit buildings. In Panel B, we restrict the sample to large building complexes. The average share of immigrants by neighborhood becomes larger at 5.8%, while the share by building complexes becomes 6.1%.

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

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 linearly. The share of Western immigrants remained stable at 2% for several years before doubling in the last decade. 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 rate of native turnover by neighborhood, which is 0.10% (standard error

[10] See Appendix B for a description on how we calculated these neighborhood amenities.

is 0.024). The positive correlation suggests that researchers must tread with caution in this type of study, as international migrants might have been moving to places that may have had high residential churn regardless.

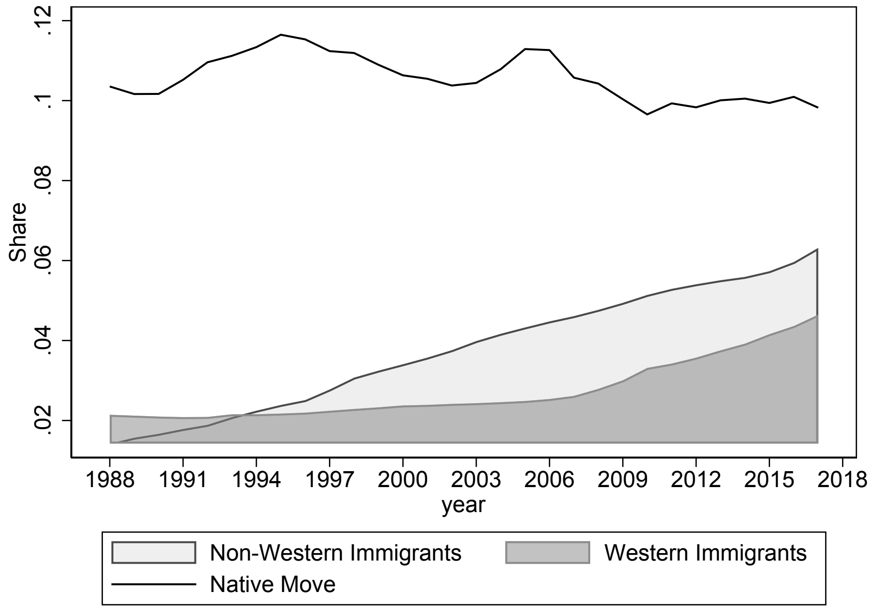


Figure 3. Long-Term Trends in Immigrant Share and native relocation, 1988-2017. This figure plots the average share of Danish residents moving to a different address each year (black line) and the shares of Western (dark-gray) and non-Western (light-gray) immigrants in the country's population.

2.4 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 B.1 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 omitting 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 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 also 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).

Housing Price Indexes. In order to create housing price indexes at the neighborhood level, we use Danish administrative register data on real estate transactions for the period 1993-2017. We restrict our focus solely to arm's length sales. These data include structural dwelling attributes extracted from the Building and Dwelling Register (BBR), such as the size of the housing unit, the number of rooms, the type of unit (multifamily house, single-family house), and the age of the building.[11] We exclude dwellings sold at prices that are above the 99th percentile and below the 1st percentile, as these likely represent outliers.

3 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 within groups. More problemat-

[11] As Table A.1 in Appendix A.5 shows, the descriptive statistics for the dwelling attributes reveal a mean price of about 1.3 million DKK (1 DKK $\approx$ 0.13 EUR.), an average size of 124 square meters, an average amount of 4.3 rooms, and an average age of 54 years.

ically, reduced *net* flows of majorities could be mechanically associated with minorities' arrivals, as the housing units occupied by the latter become unavailable (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*.[12] 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:

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

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

In addition, to control for unobservable time- or location-specific factors that affect the

[12] 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.

decision to move, we also include region-by-year fixed effects,[13] $\lambda_{r,t}$, 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. ϵ_{it} is a random error term. We cluster the standard errors at the neighborhood level in all specifications.

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 international 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.

3.1 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 number of employees—then the estimated coefficient of interest would be downward biased. On the other hand, if immigrants settle in neighborhoods with worsening job prospects, the coefficient could be upward biased.

We thus 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 A.6—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, which we also include in our specifications.

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

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, $A_{i,t}$.

$$R_{i,t} = S_{i,t} + D_{i,t} + A_{i,t} \tag{2}$$

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

$$\begin{aligned} R_{i,t} &= R^{Pb}_{i,t} + R^{Pv}_{i,t} \\ &= S^{Pb}_{i,t} + S^{Pv}_{i,t} + A^{Pb}_{i,t} + A^{Pv}_{i,t} + D^{Pb}_{i,t} + D^{Pv}_{i,t} \end{aligned} \tag{3}$$

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 $A^{Pb}_{i,t}$ 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, $\hat{\theta}^{R}_{i,1987}$, 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:

$$\hat{A}^{Pv}_{i,t} = \hat{\theta}^{R}_{i,1987} \cdot \sum_{j \neq i} A^{Pv}_{j,t} \tag{4}$$

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:

$$\begin{aligned}\hat{S}_{i,t}^{Pb} &= \left(1-\hat{\rho}_{i,t}^{Pb}\right)\cdot R_{i,t-1}^{Pb}\\ \hat{S}_{i,t}^{Pv} &= \left(1-\hat{\rho}_{i,t}^{Pv}\right)\cdot R_{i,t-1}^{Pv}\end{aligned} \tag{5}$$

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 (5) to predict the overall number of refugees moving to a different neighborhood:

$$\begin{aligned}\hat{N}_{i,t,Pb} &= \hat{\rho}_{i,t}^{Pb}\cdot R_{i,t-1}^{Pb}\\ \hat{N}_{i,t,Pv} &= \hat{\rho}_{i,t}^{Pv}\cdot R_{i,t-1}^{Pv}\end{aligned} \tag{6}$$

$\hat{N}_{i,t,Pb}$ 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:

$$\begin{aligned}\hat{M}_{i,t}^{Pb} &= \hat{\delta}_{i,t}^{Pb}\cdot \hat{N}_{i,t,Pb} + \left(1-\hat{\delta}_{i,t}^{Pv}\right)\cdot \hat{N}_{i,t,Pv}\\ \hat{M}_{i,t}^{Pv} &= \hat{\delta}_{i,t}^{Pv}\cdot \hat{N}_{i,t,Pv} + \left(1-\hat{\delta}_{i,t}^{Pb}\right)\cdot \hat{N}_{i,t,Pb}\end{aligned} \tag{7}$$

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

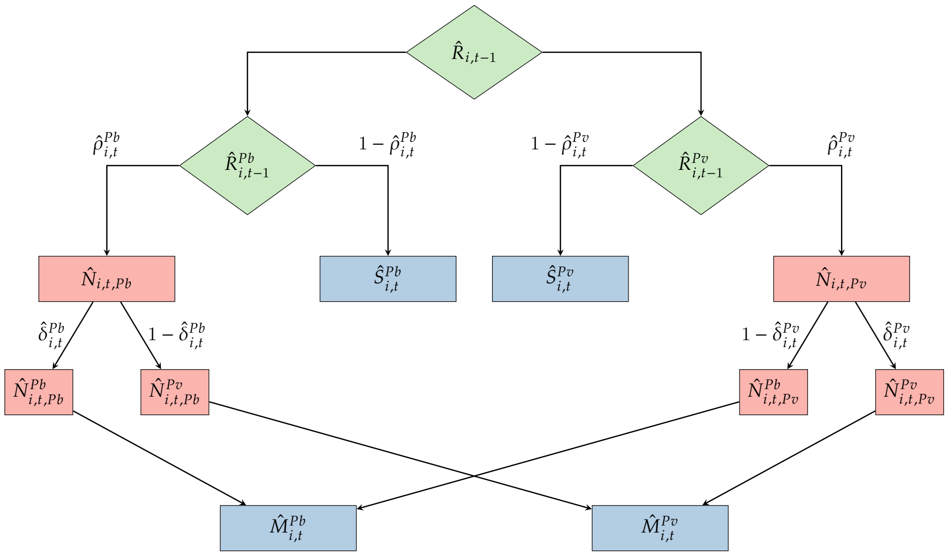


Figure 4. IV Design - Decomposing Refugees into Predicted Stayers and Movers. This figure illustrates the Markov process of refugee immigrants ($\hat{R}_{i,t-1}$) in neighborhood i, into the expected number of refugees that stayed in public housing ($\hat{S}^{Pb}_{i,t}$) or private housing ($\hat{S}^{Pv}_{i,t}$),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

$$
\begin{aligned}
\hat{D}^{Pb}_{i,t} &= \hat{\theta}^{Pb}_{i,1987} \cdot \sum_{j \neq i} \hat{M}^{Pb}_{j,t} \\
\hat{D}^{Pv}_{i,t} &= \hat{\theta}^{R}_{i,1987} \cdot \sum_{j \neq i} \hat{M}^{Pv}_{j,t}
\end{aligned}
\tag{8}
$$

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 (3) and the forecast flows discussed above, we compute the predicted number of refugees in neighborhood i and year t as follows:

$$
\begin{aligned}
\hat{R}_{i,t} &= \hat{R}^{Pb}_{i,t} + \hat{R}^{Pv}_{i,t} \\
&= \hat{S}^{Pb}_{i,t} + \hat{S}^{Pv}_{i,t} + A^{Pb}_{i,t} + \hat{A}^{Pv}_{i,t} + \hat{D}^{Pb}_{i,t} + \hat{D}^{Pv}_{i,t}
\end{aligned}
\tag{9}
$$

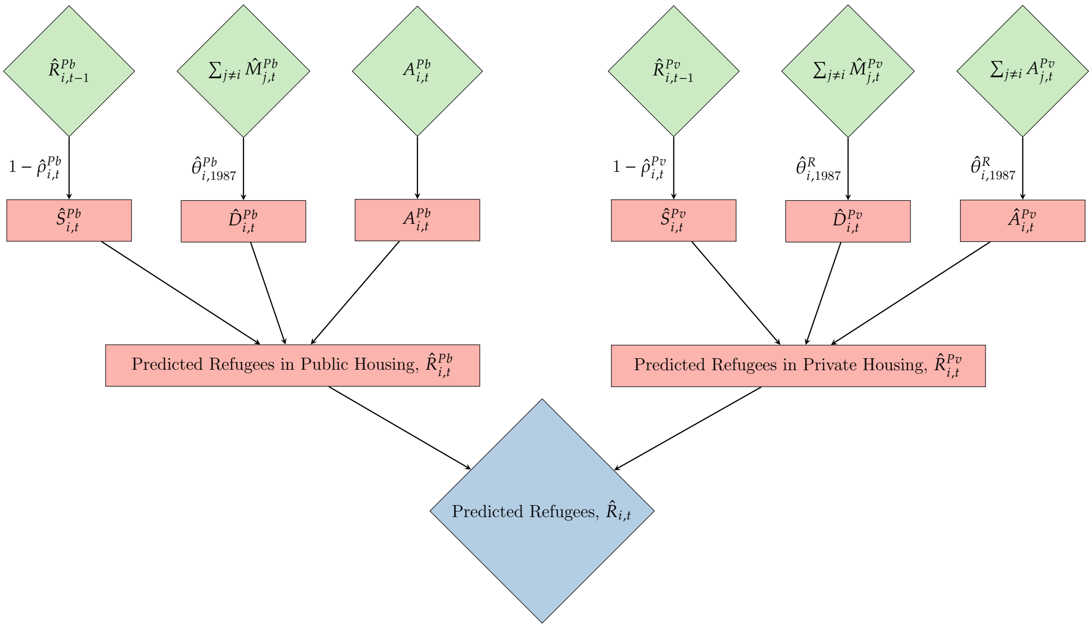


Figure 5. IV Design - Structure Summary. This figure summarizes the structure we adopt to compute the predicted number of refugees in neighborhood i, and year t.

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:

$$\text{ASM-IV}_{i,t} = \frac{\hat{R}_{i,t}}{L_{i,t}} \tag{10}$$

Of course, not every non-Western immigrant entered Denmark as a refugee. However, economic immigrant's 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 shift-share instruments based on the initial location and growth by ethnic group tend to be weak for non-refugees.

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. Nevertheless, refugees are at the center of current political debates. Therefore, this may be seen as a relevant LATE. Moreover, in practice, IV results are close to OLS ones, and we explicitly examine treatment effect heterogeneity later on.

3.2 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 over-saturated models are designed to capture differential local trends. Nonetheless, there admittedly 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 tempted to recalculate a Markov process using refugee arrivals as shocks, but randomizing their subsequent moves. However, we found that 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.[14] 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:

$$\text{ASM-IV}_{i,t} = \alpha_t + \beta_{1,t} \cdot \hat{\theta}^{Pb}_{i,1987} + \beta_{2,t} \cdot \hat{\theta}^{R}_{i,1987} + \epsilon_{i,t} \tag{11}$$

As described in Section 3.1, $\hat{\theta}^{Pb}_{i,1987}$ and $\hat{\theta}^{R}_{i,1987}$ are the public housing share and refugee share in neighborhood i in 1987, respectively. Denoting by $\hat{\alpha}_t$, $\hat{\beta}_{1,t}$, and $\hat{\beta}_{2,t}$ the estimates of α_t, $\beta_{1,t}$, $\beta_{2,t}$, we compute the following two orthogonal components:

$$\begin{aligned} \text{Shares Component}_{i,t} &= \hat{\alpha}_t + \hat{\beta}_{1,t} \cdot \hat{\theta}^{Pb}_{i,1987} + \hat{\beta}_{2,t} \cdot \hat{\theta}^{R}_{i,1987} \\ \text{Shocks Component}_{i,t} &= \hat{\alpha}_t + \hat{\beta}_{1,t} \cdot \bar{\hat{\theta}}^{Pb}_{1987} + \hat{\beta}_{2,t} \cdot \bar{\hat{\theta}}^{R}_{1987} + e_{i,t} \end{aligned} \tag{12}$$

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

[14] 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.

4 Results

4.1 Baseline Regressions

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

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: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

In column (2), we add household characteristics and time-varying neighborhood attributes. These have a substantial impact on the estimates suggesting that—indeed—the underlying mobility of natives was stronger in high-immigration locations and times. This point needs to be considered in future studies of native/white flight, which ignore composition effects driven by native heterogeneity.

The problem is parallel to the distinction between gentrification and displacement (Pennington, 2022) in research studying how the arrival of high socioeconomic status (SES) individuals affects the mobility of low-SES families. Because churn is very high among low-SES households, high-SES can progressively enter gentrifying neighborhoods without necessarily displacing any current residents, i.e., without increasing their probability of out-migration. We conclude that explicitly and empirically modeling the baseline propensity for mobility of the current native/white population, as we do here, is critical for future work.

In quantitative terms, the results still unveil substantial flight. An increase of 30 percentage points in the share of non-Western foreigners is associated with 0.63 percentage-point higher out-mobility from the neighborhood on any given year, 7.9 percent higher than the baseline. Naturally, the overall hazard of a move compounds this increased probability over the years.

In column (3), we separate the observations pertaining to 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.[15] 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.[16] 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, many interactions between people in denser areas are actually 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 more relevant phenomenon*, at the cost of sacrificing valuable variation *across* typologies.

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

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

4.2 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, 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.[17] 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 than before, but we now reject strict equality between the two instruments. This happens because the "shocks" instrument becomes rather weak by itself: it is arguably too much to ask for its variation to be strong in such an over-saturated model, which uses only 1/4th of the sample and only variance *between* large buildings. While we are confident that the instruments capture exogenous variation across neighborhoods, in the section below we use a different source of variation to convince ourselves that results are robust within the subsample of large buildings.

4.3 Neighborhood and year Fixed Effects

As in past research, so far we have kept with conventional neighborhood estimates. However, having an immigrant immediately next door may be as important for native behavior. We now go back to Table 3, column (6), and leverage the richness of the data to investigate additional impacts at very local interaction levels. Concretely, we focus on

[17] In Table C.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 C.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.

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 3.1. Panel B shows the results when we use the orthogonal shares and shocks components as described in Section 3.2.

616,094 residential structures containing more than 10 apartment or condo units.[18].

To estimate effects above and beyond those captured by immigrant neighborhood shares, we now deploy both neighborhood-by-year and building-complex fixed effects. An advantage of saturating the model so aggressively is that the neighborhood-by-year fixed effects should capture any other conceivable time-varying unobservable shocks that impacted native mobility.

Note that this specification utilizes 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. While the definition of the treatment is not the same, finding native flight here should add substantial credibility to the existence of this phenomenon.

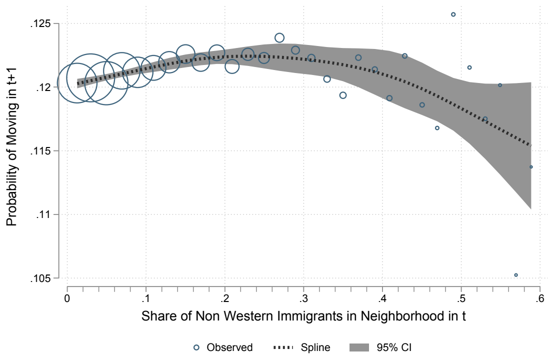
Despite soaking up a large amount of the variation in the data, we still find *additional* effects of the micro concentration of non-Westerners on native exits across buildings in the same area and year. For instance, consider two housing complexes within the same neighborhood: one with a 30 percent foreign-born share versus another with an all-indigenous population. On average, we will see 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. Ethnic interactions thus appear to layer up at different geographic extents.

4.4 Non-linearities

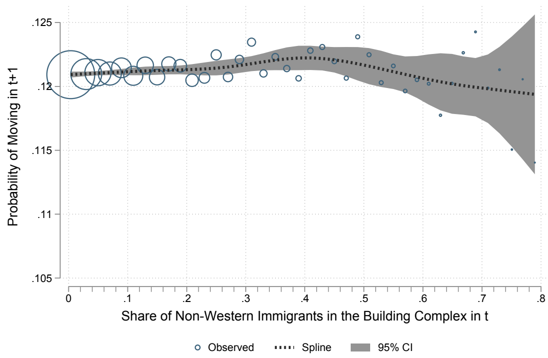
Our baseline specifications assume a linear relationship between the growth of the immigrant share and the probability of native exits. However, the propensity of locally-born citizens to move out could increase in locales with higher baseline concentrations of the foreign-born (Schelling, 1971; Card et al., 2008). In this section, we assess whether native flight is nonlinear. We restrict the analysis 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 (1), including the set of fixed effects, but excluding immigrant shares. We then compute the adjusted probability of moving out by adding back the residuals to the regression-predicted probability of exit *at national average characteristics*. We subsequently calculate the local averages of such adjusted probabilities—by neighborhood or building complex–grouped into bins according to their lagged shares of immigrants. We finally estimate a continuous and smooth restricted cubic spline that models the local means of the compositionally-adjusted probability of moving out as a function of aver-

[18]Note that immigrant shares 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 C.2 of Appendix C.2)



(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.

age immigrant density in the previous year. We weigh the estimates by the number of locations—neighborhoods or building complexes—within each bin.

This specification is—to the best of our knowledge—novel to the literature. Previous work primarily focuses on net outflow-to-inflow relationships, which will miss turning points in native exits that are compensated by countervailing native arrivals. In other words, conventional tipping point estimates ignore processes of spatial sorting of majorities by distaste for minorities. In our approach, finding a tipping point is a sufficient condition to show that the local distribution of native tastes changes nonlinearly.[19]

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

Panel A shows a mildly inverted U-shape relationship between the share of immigrants and the probability of moving: the likelihood of a native family moving out of its neighborhood increases with immigrant density up to around 30%, but then starts decreasing to reach baseline back at 43%. This seems consistent with the existence of a distribution of native tolerance (or moving costs): less accommodating households (or those with low mobility costs) flee first until the marginal native is indifferent to the composition of the neighborhood (or displays high mobility costs).

In other words, at about 43% non-Westerner share, everyone who wanted to disproportionately move out has already done so. At that point, further neighborhood ethnic

[19]However, it is not a necessary condition, as the majority population *inflows* might still depend nonlinearly on the immigrant share, even if *outflows* did not.

change may still happen, but it must now go via replacements in naturally-occurring vacancies. Nevertheless, confidence bands also increase substantially at high foreign-born shares, so any conclusions here must be tentative. In any case, there is no clear tipping point.

In Panel B, we observe a similar pattern for the share of immigrants in the building complex, with the turning point at around 40%. However, this extremely-saturated model also becomes imprecise at high rates of exposure. We conclude that there is no evidence of tipping point dynamics, whereby native exits accelerate after a threshold is crossed.

4.5 Heterogeneity of Native's Responses

We next study treatment effect heterogeneity: which types of natives are more prone to flight? 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. Concretely, we estimate the effect of the neighborhood share of immigrants as in Table 3, column 5, but now interacted with dummies for each group:

$$\begin{aligned} P(y_{i,z,t+1} = 1 | \mathbf{X}_{i,z,t}, s_{z,t}, \lambda_{r,t}, \lambda_z) =& \alpha + \Sigma_{g=1}^{k} \beta_g \cdot s_{z,t} \cdot \mathbb{1}(G_i = g) \\ & + \Sigma_{g=1}^{k} \mathbb{1}(G_i = g) + \Theta' \mathbf{X}_{i,z,t} + \lambda_{r,t} + \lambda_z + \epsilon_{i,z,t+1} \end{aligned} \tag{13}$$

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, are restricted to building complexes with more than ten units. Standard errors are clustered at the neighborhood level. In order to avoid type I errors or publication biases, Section 4.5.4 shows that significant results can be reproduced in alternate random split samples.

4.5.1 Heterogeneity by Age, Income, and Education

We begin by grouping natives according to their age, education, and income levels. Household heads with ages below the median (47 years old) are classified as young, or old otherwise. We also classify households as low or high-income, depending on whether their disposable household income was below or above the within-year median. Individuals are classified as having low education if their highest attainment was primary school (38 percent of the sample), and as having high education if they completed vocational training or further. Figure 7 plots the estimated coefficients for the interaction terms. We find

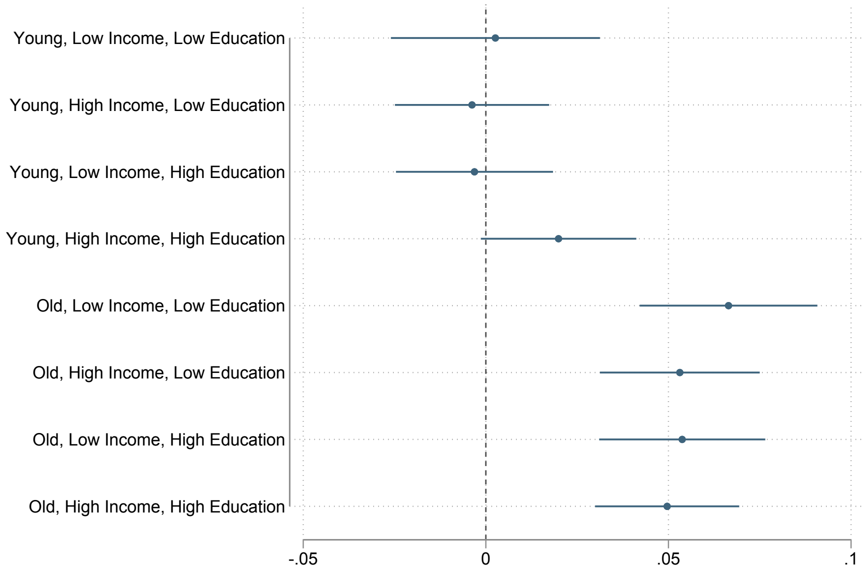


Figure 7. Heterogeneity by Age, Education, and Income. In this figure, we plot the estimated coefficients for the heterogeneity analysis as described in Section 4.5.1. The coefficients are the interaction terms of the share of non-Western immigrants in the neighborhood with binary variables corresponding 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 older heads of household tend to move more in response to non-Western immigrants and that this response is very similar across different levels of education and income. We do not find statistically significant effects for young heads of households, regardless of their income and education levels.

4.5.2 Treatment Effects by Age, Income, and Parental Status

We now consider that households with children might face distinctive constraints or preferences for local amenities. In the US, Cascio and Lewis (2012) find that parents' school choices may drive the ethnic segregation of foreign-born children. We estimate a specification as in the previous section but replace educational levels with a binary variable assuming a value of 1 for households with children below 18 years old.

Figure 8 plots the coefficients, which reinforce that older heads of households are more likely to exit. In contrast, the presence of children in the household does not seem to be associated with increased flight responses. One potential explanation for this result is the relocation of children to other public or private schools—rather than moving out the whole family to a different neighborhood (Farre et al., 2018; Bjerre-Nielsen and Gandil, 2024).

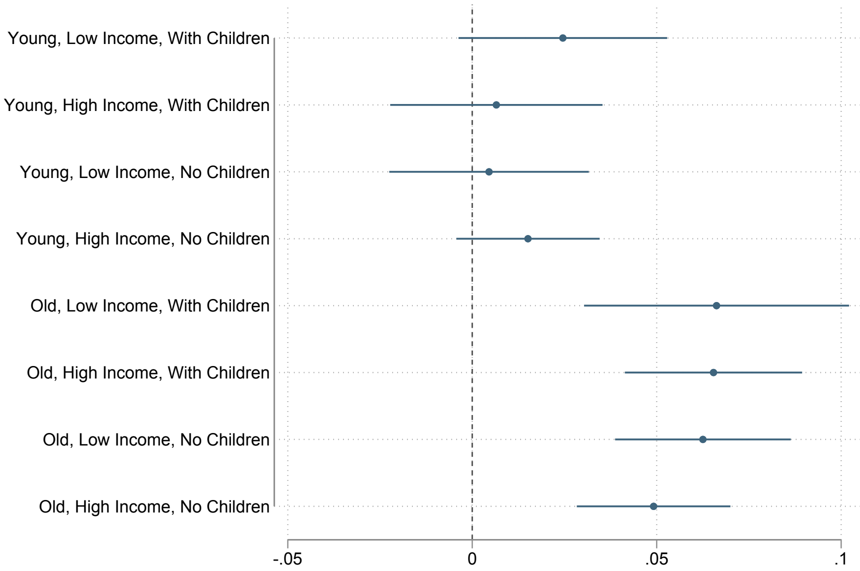


Figure 8. Heterogeneity by Age, Income, and Children in the Household. In this figure, we plot the estimated coefficients for the heterogeneity analysis as described in Section 4.5.2. The coefficients are the interaction terms of the share of non-Western immigrants in the neighborhood with binary variables 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.

4.5.3 Treatment Effect Heterogeneity by Housing Tenure

Exiting research indicates that homeowners move less often (Jia et al., 2023). Migration costs may, therefore, mediate reactions to local ethnic shocks. Another important characteristic is whether families live in social housing, representing a significant portion of the stock in Denmark: 42% of rental units (Figure A.4). Research in both Denmark (Munch and Svarer, 2002) and the Netherlands (De Graaff et al., 2009) has established a lock-in effect: moving out of social housing is associated with substantial costs for households, as rents are below market ones, and the quality of the units is generally good.

Current social housing renters in Denmark obtain priority in the waiting lists for other units administered by the same housing organization—especially upon divorce or childbearing. However, waiting lists in the major cities are years-long regardless: in April 2024, 1,095 units—mostly in faraway areas—were available compared to 601,862 people in waiting lists in the whole country.[20] Therefore, it is difficult to swap a social unit for another one.

We thus estimate Equation (13) defining groups as (i) renters living in private housing, (ii) renters living in social housing, or (iii) homeowners. Figure 9 plots the estimated

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

coefficients for the interaction terms, β_g.

We find a negative coefficient for renters in private apartments. However, this estimate does not survive our split sample robustness test below and likely arose by chance. The coefficient for condo homeowners is not statistically significant. However, a robust finding is that of 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 salient public housing minority "ghettos." While public perceptions are not always empirically right, and may be exaggerated, our evidence here is consistent with the existence of ethnic tensions in the social housing sector. However, per these results, any issues with regard to ethnic segregation in social housing have at least as much to do with native aversion as they are products of the tendency for low-income, foreign-born families to cluster together. This suggests that policy approaches cannot simply focus efforts and rhetoric on the latter but must also confront and address the former.

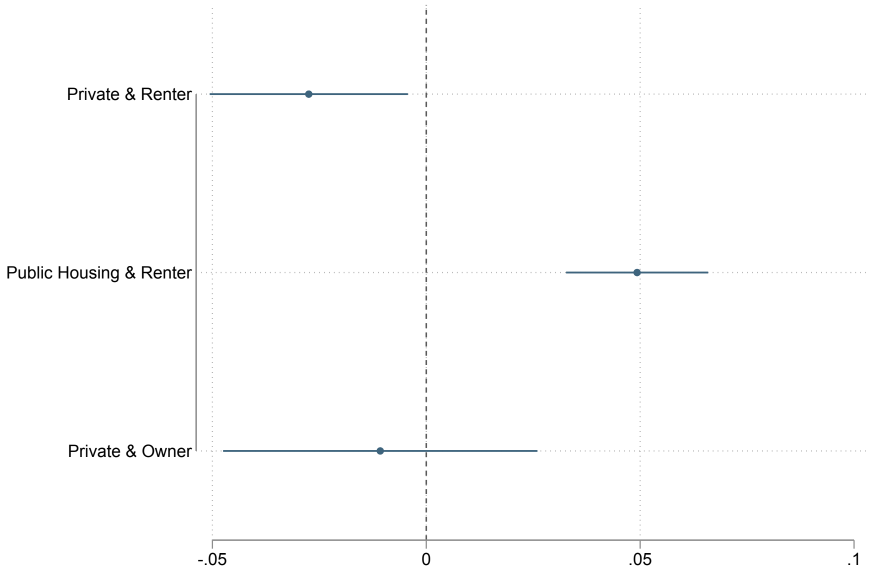


Figure 9. Heterogeneity by Housing Tenure. This figure compares the effects of the immigrant presence on the propensity to move for three groups: public housing renters, private unit renters, and private unit homeowners. The dots are the estimates for the interaction terms of the share of non-Western 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.

4.5.4 Multiple Testing

Because we have performed several hypothesis tests, a natural concern is that some of the observed significant results represent false rejections of the null. Conventional Bonferroni-style adjustments can be performed in these situations.

However, since we are blessed with an abundance of data, it is more convincing to randomly split the sample into two sets of neighborhoods and run the exact same set of regressions on each set separately (Anderson and Magruder, 2017).[21] Conventional adjustments (Benjamini et al., 2006; Romano and Wolf, 2005) are based on joint hypotheses encompassing all parameters.[22] Unappealingly, such adjustments depend on the *reported* number of coefficients estimated by the researcher. Split sample approaches validate specifically the parameter of interest more efficiently and are independent of the purported number of parameters being tested. By replicating the results in two independent and random samples, we reduce the probability of type I errors under the null to 0.0025.

Figures C.1, C.2 and C.3 in Appendix C.4 show the estimated coefficients for the share of immigrants and their interaction terms with each of the groups described earlier, but for two random partitions of the sample. The results confirm that native flight comes mainly from older household heads, whereas income, education, and the presence of children do not seem to matter. The response of private renters is the only non-robust coefficient, as it does not replicate throughout both samples.

5 Heterogeneity by Immigrant Type

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

In this section, we assess the potential heterogeneity of the estimates to differences in the characteristics of the immigrants. First, we group immigrants according to their household income, with the ones above the median being classified as high-income, or low-income otherwise. We further split the immigrants according to four different cultural and socioeconomic characteristics. To test the role of religion, we define a binary variable

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

[22]For instance, they may consider a 5 percent confidence for the probability that *any* of the estimators in a group suffers from type I error. Because we generally do not know *which* of the individual parameters was wrongly deemed as significant in the separate regressions, such adjustments need to be very conservative on all. More so the more parameters we estimate. Therefore, type II errors on each individual parameter can be larger than desirable, especially as the number of tests increases.

that equals one if the immigrant is from a country where the majority of the population is Muslim and zero otherwise. We allocate immigrants into a high language proximity group if they hail from countries with a similarity index to Danish above the median or into a low language proximity group otherwise. High social connectedness is assigned to foreigners from countries displaying Facebook connectedness indexes above the median vis-a-vis Denmark (or low social connectedness otherwise). Finally, we use the World Bank country income classification groups to separate low or middle-low-income countries of origin from middle-high or high-income ones. We then estimate the following specifications:

$$P(y_{i,z,t+1} = 1|\mathbf{X}_{i,t}, s_{z,t}, \lambda_{r,t}, \lambda_z) = \alpha + \Sigma_{g=1}^{4}\beta_g \cdot s_{g,z,t} + \Theta'\mathbf{X}_{i,t} + \lambda_{r,t} + \lambda_z + \epsilon_{i,z,t+1} \quad (14)$$

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 3. In each of four separate regressions, we group immigrants into another four categories, using household income and one of the measures described above. In order to focus on places 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 shows parameters estimates and confidence bands for Muslim migrants. The presence of low-income immigrants from Muslim countries has a positive and statistically significant effect on native exits, but high-income Muslim immigrants have a negative impact. A one percentage point increase in the share of low-income immigrants from Muslim countries in the neighborhood (building complex) increases the probability of a native moving out by 0.13 percentage points on average.

With regard to social proximity, low-income immigrants are associated with flight regardless of language proximity. The estimates for groups based on language similarity or the wealth of the country of origin repeat this pattern. The evidence shows that—rather than based on religion, language, culture, or national origin—native flight in Denmark seems to be driven by avoidance of foreign-born individuals with low socioeconomic status. These results are consistent across random split samples (as can be seen in Figure C.4 in Appendix C.4).

6 Incoming Residents and Native Sorting

Mobility responses to ethnic shocks are sufficient to establish the existence of in-group preferences. However, long-term neighborhood outcomes also depend on who is moving in to fill vacancies as they arise.

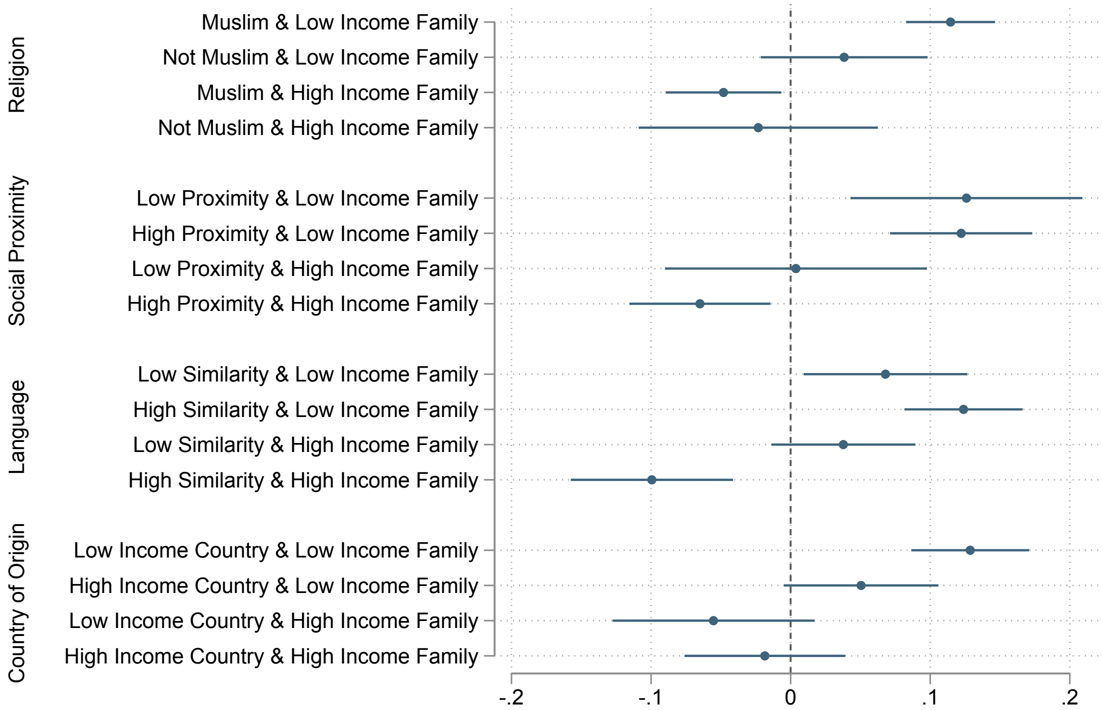


Figure 10. Heterogeneity by Immigrant Group. In this figure, we plot the estimated coefficients for the heterogeneity analysis as described in Section 5. We estimate four regressions where we separate non-Western immigrants based on their household income level and into categories based on their country 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.

6.1 Immigrant Snowballing

We begin by investigating potential foreign-born snowball patterns of settlement. We now condition on a subsample that only includes vacant housing units. Because our data are annual, units are considered vacant for some period if we find the reference householder at time t living in a different unit at $t+1$. We can then categorize the identity of the incoming head of such vacant unit at $t+1$, or as soon as the unit becomes occupied again. We want to learn about who is moving into areas with high minority concentration, and to do so, we estimate the following linear probability model:

$$P(M_{i,z,t+1} = 1 | y_{i,z,t+1} = 1, \mathbf{X}_{i,z,t}, s_{z,t}, \lambda_{r,t}, \lambda_z) = \alpha + \beta \cdot s_{z,t} + \Theta' \mathbf{X}_{i,z,t} + \lambda_{r,t} + \lambda_z + \epsilon_{i,z,t+1} \quad (15)$$

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 share of minorities in the neighborhood increases the probability of a non-Western immigrant moving into a vacant unit by 5 percentage points. Similar results are obtained for multifamily buildings (column 5). This is a very large effect compared to a national average of 6 percent of non-Westerners in large buildings.

Nonetheless, these results alone do not necessarily imply native avoidance: they could simply reflect a revealed preference for clustering among immigrants. Under such a scenario, non-Westerners could be outbidding natives for their preferred locations close to other migrants. However, we later show evidence on housing prices that casts doubt on such an interpretation.

Column (6) controls again for neighborhood and year-fixed effects and focuses on the share of immigrants by building a complex. Remember that these effects are additive to the ones at the neighborhood level. We still find disproportionate immigrant clustering above and beyond the latter. A building with an extra 10 percent of non-Westerners would see an additional probability of immigrant entry of about 2 percentage points.

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***	0.502***	0.471***	0.291***	0.514***	
	(0.022)	(0.022)	(0.052)	(0.023)	(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: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

6.2 Sorting by Native Avoidance

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

To investigate this issue, we now condition on the subsample of non-immigrant families moving into a new home during the data period. We run similar specifications as in equation (15), where the dependent variables are four indicators for whether the new resident householder 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 4.5.

Table 6, shows that higher local shares of non-westerners are associated with the arrival of Danish-born families that are disproportionally young, low-income, and without children. We do not find statistically significant coefficients on educational levels.

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: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Hence, the results are similar to those pertaining to flight regarding the age of migrant avoiders. However, the presence of children is now very strongly associated with native avoidance. This result can be rationalized through the high cost of removing children from their social environment. Once a family decides to move—and moving costs are already sunk—they prefer to avoid migrant-dense neighborhoods.

These results may also be partially driven by the structure of social housing benefits in Denmark (20 percent of the stock): after a family with children has already moved into a social home, it is very costly to give up the unit and go back to the waiting list. However, it is trivial to simply select a more desirable set of housing complexes at the time of enrollment into the waiting list. The associations between social housing and immigrant arrivals have been explored in France by Verdugo (2016); Verdugo and Toma (2018), but future research in this area—including native behavior and native-migrant relationships within these complexes—is needed.

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

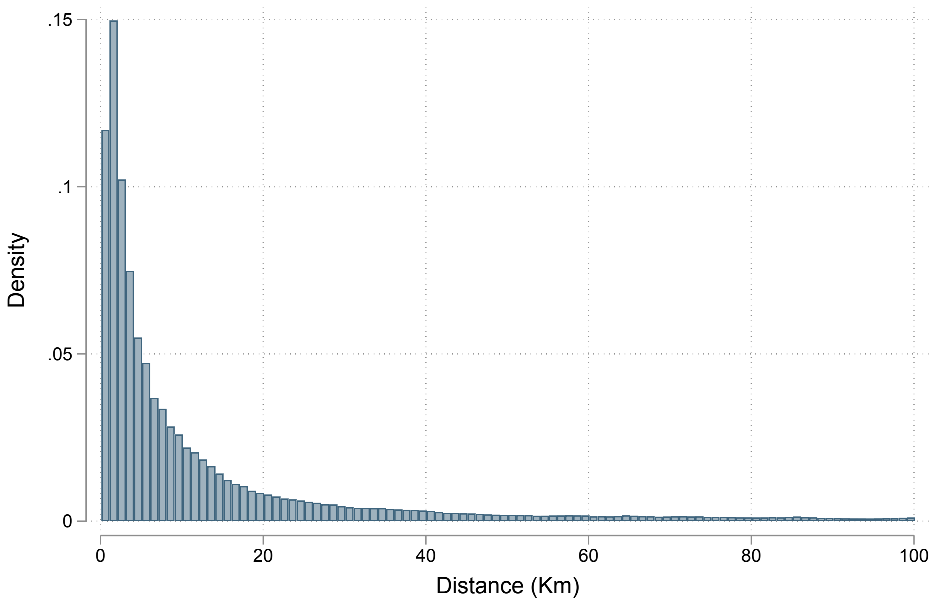


Figure 11. Histogram of Distances Between Origin and Destination Neighborhoods. In this figure, we plot the histogram of distances in kilometers between the origin and the destination neighborhoods for all movers in our data with distances below 100 kilometers. Figure A.5 in the Appendix shows the complete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

7 Flight to Where?

We now focus on the ethnic characteristics of the destinations for families moving away from minority areas. The major challenge in that regard originates 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 moving at random, we should still measure lower foreign-born shares in destination neighborhoods.

However, people do not move at random. In fact, moving distances are characterized by a strong gravitational pattern. Conditioning on moves below 100km for visibility, Figure 11 plots the histogram of distances between origin and destination neighborhoods for all movers in our data. When people change residences, they tend to disproportionately move 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 weights are based on a gravity equation:

$$\text{RWAIS}_{z,t} = \sum_{j \neq z}^{N} \frac{s_{j,t}}{d_{jz}^{\beta} \left(\sum_{j \neq z}^{N} \frac{1}{d_{jz}^{\beta}} \right)} \tag{16}$$

where $s_{j,t}$ 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 ($\beta = 2$).[23] 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 relevant sample of Danish natives who were living in large building complexes and moved to a different neighborhood in the following year.[24] 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, slope of the blue line indicates a positive relationship between the immigrant shares at the origin and destination. Second, the slope of the curve becomes flatter as we increase the share of immigrants. However, these basic patterns are very simply explained by the distribution of immigrant shares in proximate neighborhoods, as captured by the counterfactual—RWAIS—red line. Naturally, people living in the central areas of Copenhagen who need to move to similar neighborhoods in the city will be exposed to higher immigration shares than in the rest of the country. As pointed out by Blair (2023), outside options are important.

Nevertheless, our third takeaway is that the actual exposures for locals in destination neighborhoods are below their counterfactual. Moreover, the gap between observed and predicted immigrant shares at the destination increases as the share at the origin grows higher. Therefore, we conclude that natives moving out of immigrant-dense neighborhoods are looking for new homes in areas that are distinctively more "native."[25]

[23]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, obviously very close to our baseline choice.

[24]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.

[25]Appendix C.5 shows that the main takeaways discussed here are robust to different values for the spatial decay parameter, β.

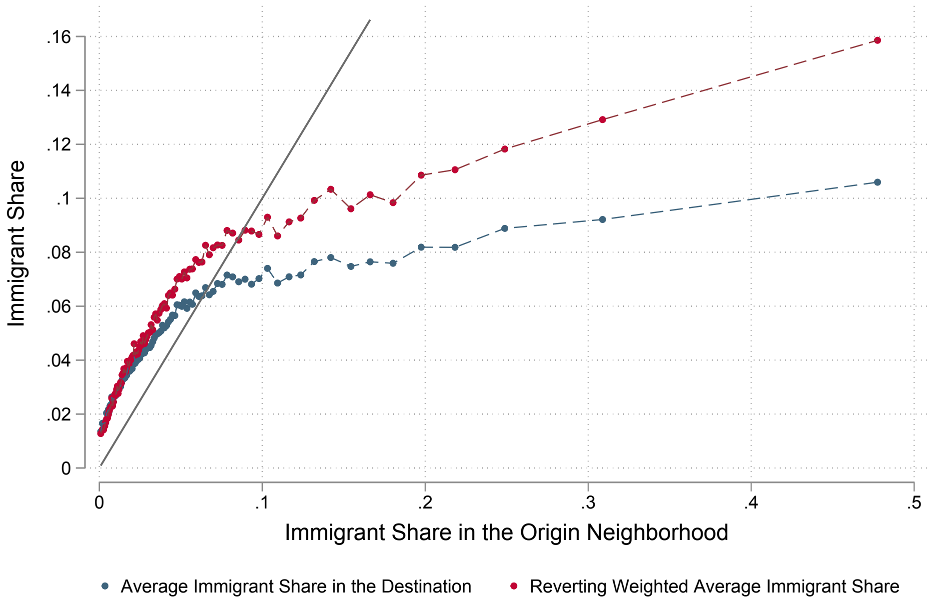


Figure 12. Association between origin's and destination's immigrant shares. This figure shows the relationship between the immigrant shares at the origin and destination for all Danish citizens and heads of households who moved to a different neighborhood. It plots the observed mean immigrant share at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

8 Housing Prices

Empirical studies comparing metropolitan areas tend to find positive impacts of immigration on house prices due to increasing populations in destination cities (Cochrane and Poot, 2021).[26] However, *within* metropolitan areas, most studies find that concentrations of the foreign-born and minorities have a deflating effect on housing values at the neighborhood level.[27] These results are consistent with *native-flight* and sorting by ethnicity, with natives displaying a higher willingness-to-pay (WTP) premium to live in more homogeneous neighborhoods.

To assess the issue in the current context, we use housing transaction data spanning most of our target period (1992-2017). This period encompasses different stages of the housing market boom-bust cycle, as well as institutional changes.[28] We start by

[26]For example, Saiz (2003, 2007), Ottaviano and Peri (2012) and Sharpe (2019) provide estimates for the US, Akbari and Aydede (2012) for Canada, Sanchis-Guarner (2023) for Spain and Moallemi and Melser (2020) for Australia. On the other hand, Cortes and Sant'Anna (2023) provides estimates for the case of an outflow of immigrants in the US.

[27]See for example, Hatton and Tani (2005); Saiz and Wachter (2011); Accetturo et al. (2014); Sá (2014). There is an extensive earlier literature on housing price declines in neighborhoods that experienced African American population arrivals in the US (e.g., Berry (1976); Coulson and Bond (1990)).

[28]The interested reader should consult Appendix A.5 for details. For instance, Figure A.6 in Appendix A.5 illustrates the changes in the number of transactions over this period. Notably, a tremendous

estimating the following model:

$$p_{h,t,z} = \alpha + \beta \cdot s_{t-1,z} + \Theta' \mathbf{X}_{i,t} + \lambda_z + \rho_{t\times M(h)} + \epsilon_{i,t,z} \tag{17}$$

where $p_{h,t,z}$ is the log of house price of house h in year t in neighborhood z; $s_{t-1,z}$ is the lagged non-Western immigrant share in z; $\mathbf{X}_{i,t}$ denotes structural dwelling attributes; λ_z are neighborhood fixed effects; $\rho_{t\times M(h)}$ are municipality-specific ($z \in M$) year fixed effects, and ϵ_{it} is a random error term. Neighborhood fixed effects control for time-invariant local heterogeneity, while municipality-and-year dummies control for time-varying unobserved shocks affecting each city/town, including changes to housing supply or differential trends between high- and low-density areas.[29]

The first column in Table 7 reports OLS estimates of equation (17). The coefficient on the non-Western immigrant share is negative and statistically significant. It implies that a 30 percentage point increase in the non-Western share is associated with 5.1% lower prices, corresponding to an approximate average decline of 66,000 DKK (about 8,600 EUR).[30]

Table 7. Housing prices and non-Western immigrants

	log p	log HPI	log HPI
	OLS	OLS	2SLS
	(1)	(2)	(3)
Non-western immigrant share (t-1)	–0.170***	–0.413***	–0.346***
	(0.067)	(0.013)	(0.023)
First stage results			
ASM-IV			1.789***
			(0.005)
First stage F-statistic			31,480
N	1,166,196	64,176	64,176
R^2	0.790	0.981	
Controls (housing characteristics)	✓		
Neighborhood FEs	✓	✓	✓
Year-municipality FEs	✓	✓	✓

Note: Controls (housing characteristics) include space (log sqm), age of the building and age squared, number of rooms, and type of unit (multifamily house, single-family house); standard errors clustered at the neighborhood level are in parentheses; *p<0.1; **p<0.05; ***p<0.01.

Unfortunately, however, our transaction data is sparse, with few or missing observations in several neighborhoods and years. Because these occurrences may not be random, we computed a housing price index (HPI) using all microdata and optimal imputations based on tax-assessment data (see Appendix B.2). Using an HPI has the additional ad-

decrease was observed in 2008 due to the financial crisis.

[29]Denmark is divided into 98 municipalities.

[30]The mean house price during this period is 1.307 DKK Million (about 170,000 euro), see Table A.1 in Appendix A.5.

vantage of providing complete data for the period 1987-2017, an additional five years of information.[31] The main parameter using our novel HPI for all neighborhoods and years is shown in column (2). Estimated negative coefficients on non-Western immigrants are now noticeably larger.[32]

Immigrants may choose to settle in areas with increasing economic prospects and consequently growing housing prices, or in neighborhoods that experience negative shocks and are becoming more affordable. This can lead to omitted-variable or endogeneity biases. In order to mitigate them, we deploy our ASM-IV strategy once again. Column (3) shows that the instrument remains strong, even after including the additional year-municipality fixed effects. The estimated coefficient is -0.346. It implies that a 30 percent increase in the non-Western immigrant share results in a 10% decrease in housing prices.

Overall, the results here do not support an explanation for native avoidance based on tightness in the housing market. Such an explanation would imply that non-Westerners are competing for specific locations by outbidding natives, a counterfactual claim.

9 Conclusion

Immigration has become the most salient and divisive issue in the political discourse of most European countries. 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 European 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. This literature originally focused on the segregation of the African-American population as an expression of generalized white racism. However, it later extended into the study of the endogenous dynamics that have arisen from that legacy,

[31] For the period 1987-1992, the evolution of tax assessments on each dwelling and estimates for the year 1993 are used to compute hedonically-adjusted housing price indices for each of the neighborhoods.

[32] This is likely due to the negative correlation between the share of non-Western immigrants and the number of transactions at the neighborhood level, see Figure A.7 in Appendix A.5.

[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".

and broadened its scope to immigration and the Hispanic population. While similar issues have been increasingly and adroitly studied in a European context, the depth and width of the current literature are still far from commensurate to their importance.

Here, we exploited a comprehensive, geolocated individual-property panel dataset encompassing the entire population of Denmark from 1987 to 2017. This period witnessed a substantial influx of non-Western immigrants and, in the later years, a noticeable discontent among natives regarding their rising numbers. We conclusively show that many individuals opted to express their dissatisfaction by voting with their feet, disproportionately relocating away from neighborhoods where immigrants had settled.

The data quality allows us to be exhaustive in findings and careful in assigning causality. Departing from the previous literature, we are able to model individual household decisions. This, in turn, allows us to control for factors that may be associated with high native churn, which may have facilitated immigrant access. We can also use the random nature of refugee placements to generate suitable quasi-experimental variation. In a very demanding specification, we control for neighborhood-and-year fixed effects. We can still identify native flight from buildings that hosted more non-Westerners compared to others in the same area and year.

To the best of our knowledge, we are the first to study tipping points with regard to individual "white" flight from a neighborhood—as opposed to the net overall changes in ethnic shares. This is important because flight from less tolerant natives may be compensated in the aggregate by arrivals from more accommodating ones. In doing so, we identify an intriguing inverted U-shaped relationship between the immigrant concentration and the probability of relocation. As the concentration of immigrants rises, the likelihood of native individuals moving out increases steadily, reaching its peak at around 30%. However, beyond this threshold, this probability starts to decrease, becoming similar to its baseline at about 40%. We interpret this as a sign of heterogeneous preferences: everyone who wanted to leave has already done so by then.

Our study offers other compelling evidence of heterogeneous behavior. Flight is stronger across multifamily buildings, as it is among older individuals. Moreover, our analysis reveals that the impact of immigrant presence varies depending on the specific characteristics of the immigrants themselves. Notably, flight only arises in the presence of low-income immigrants. This highlights the importance of considering the nuanced effects of different immigrant profiles when examining the dynamics of native mobility.

Areas with a higher immigrant concentration are more likely to attract new residents who are either other non-western migrants or Danish-born citizens who are young, low-income, and without children. This sorting process of natives in minority neighborhoods may require some attention from policymakers, as the families moving in might not be the best equipped to manage potential inter-community conflict.

We also find that families who move out tend to choose destinations with lower immigrant shares than predicted by a gravity relationship. Finally, our empirical results reveal the negative impacts of non-Westerner concentrations on housing prices. This allows us to negate explanations for native avoidance based on housing market tightness.

Our findings are relevant for policymakers and academics concerned with immigration and ethnic segregation. However, we acknowledge that their generalizability to alternative contexts remains uncertain. Nevertheless, the results illustrate how social interactions influenced by native preferences can hinder neighborhood integration. Furthermore, it is also essential to consider the composition of the neighborhood's native population. As a result, conventional policy instruments targeting specific geographic locations and focusing on immigrant families alone may have limited effectiveness in 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., Stephan Heblich, and Tobias Seidel, "Micro-geographic property price and rent indices," *Regional Science and Urban Economics*, 2023, *98*, 103836.

Akbari, Ather H. and Yigit Aydede, "Effects of immigration on house prices in Canada," *Applied Economics*, 2012, *44* (13), 1645–1658.

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.

_ , **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," *NBER Working Paper n.30487*, 2022.

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.

Berry, Brian JL, "Ghetto expansion and single-family housing prices: Chicago, 1968–1972," *Journal of Urban Economics*, 1976, *3* (4), 397–423.

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.

Cochrane, William and Jacques Poot, "Effects of immigration on local housing markets," in Karima Kourtit, Bruce Newbold, Peter Nijkamp, and Mark Partridge, eds., *The Economic Geography of Cross-Border Migration*, Cham: Springer International Publishing, 2021, pp. 269–292.

Cortes, Gustavo and Vinicios Sant'Anna, "Send them back? The real estate consequences of repatriations," *MIT Center for Real Estate Research Paper*, 2023.

Coulson, N Edward and Eric W Bond, "A hedonic approach to residential succession," *The Review of Economics and Statistics*, 1990, pp. 433–444.

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.

_ 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.

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.

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.

Hatton, Timothy J. and Massimiliano Tani, "Immigration and inter-regional mobility in the UK, 1982–2000," *The Economic Journal*, 11 2005, *115* (507), F342–F358.

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.

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

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

Moallemi, Morteza and Daniel Melser, "The impact of immigration on housing prices in Australia," *Papers in Regional Science*, 2020, *99* (3), 773–786.

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.

Ottaviano, Gianmarco I.P. and Giovanni Peri, "Chapter 3: The effects of immigration on US wages and rents: a general equilibrium approach," in Karima Kourtit, Bruce Newbold, Peter Nijkamp, and Mark Partridge, eds., *Migration impact assessment: New horizons*, Cheltenham, UK: Edward Elgar Publishing, 2012, pp. 107–146.

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

_ , "The spatial impact of new housing on rents, displacement, and gentrification: Evidence from San Francisco," *Working Paper*, 2022.

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.

Sanchis-Guarner, Rosa, "Decomposing the impact of immigration on house prices," *Regional Science and Urban Economics*, 2023, *100*, 103893.

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

Sharpe, Jamie, "Re-evaluating the impact of immigration on the U.S. rental housing market," *Journal of Urban Economics*, 2019, *111*, 14–34.

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.

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

_ , "Dansk Branchekode 2007," 2007.

Sá, Filipa, "Immigration and house prices in the UK," *The Economic Journal*, 09 2014, *125* (587), 1393–1424.

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

(not for publication)

A Main Dataset and Additional Descriptive Stats

A.1 Refugees in Denmark

Figure A.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.

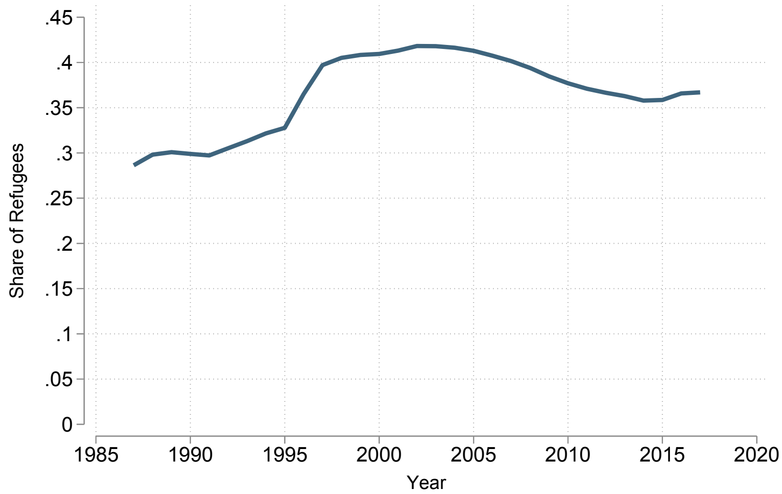


Figure A.1. Evolution of the Share of Refugees in Denmark. This figure illustrates the evolution of the share of refugees in Denmark in the total number of non-Western immigrants.

A.2 Additional Maps

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

Figure A.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).

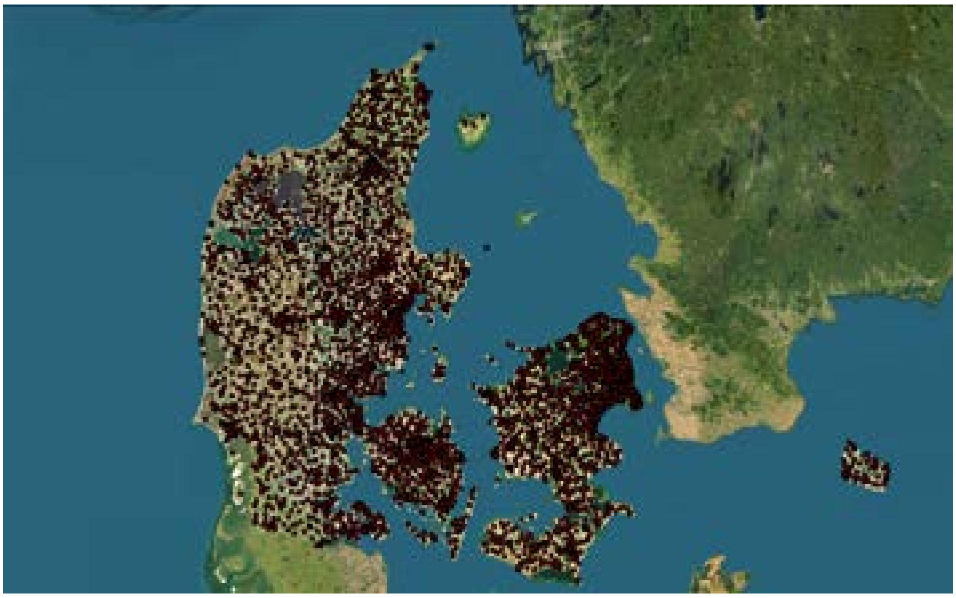


Figure A.3. Small Neighborhood Centroids. This figure illustrates the geographic distribution of small neighborhoods across Denmark. The red points are the centroids of each area, which are based on the definitions of small neighborhoods as constructed by Damm and Schultz-Nielsen (2008).

A.3 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 A.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.

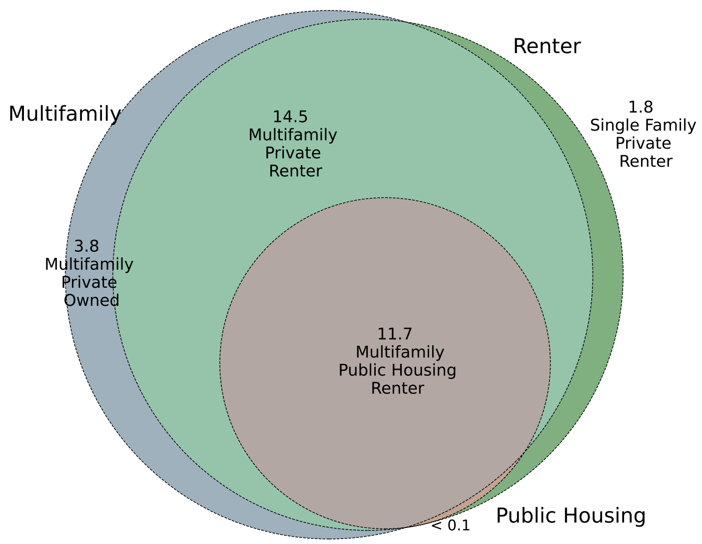


Figure A.4. Overlapping of Characteristics: Public Housing, Renters, and Multifamily Buildings. In this figure, we plot the Venn diagram to illustrate the relationship between living in multifamily buildings, in social housing, and being a renter. The numbers are in millions of heads of households in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

A.4 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 A.5, we present the complete distribution of distances. Again, we find a strong gravitational relationship that characterizes the moving distances.

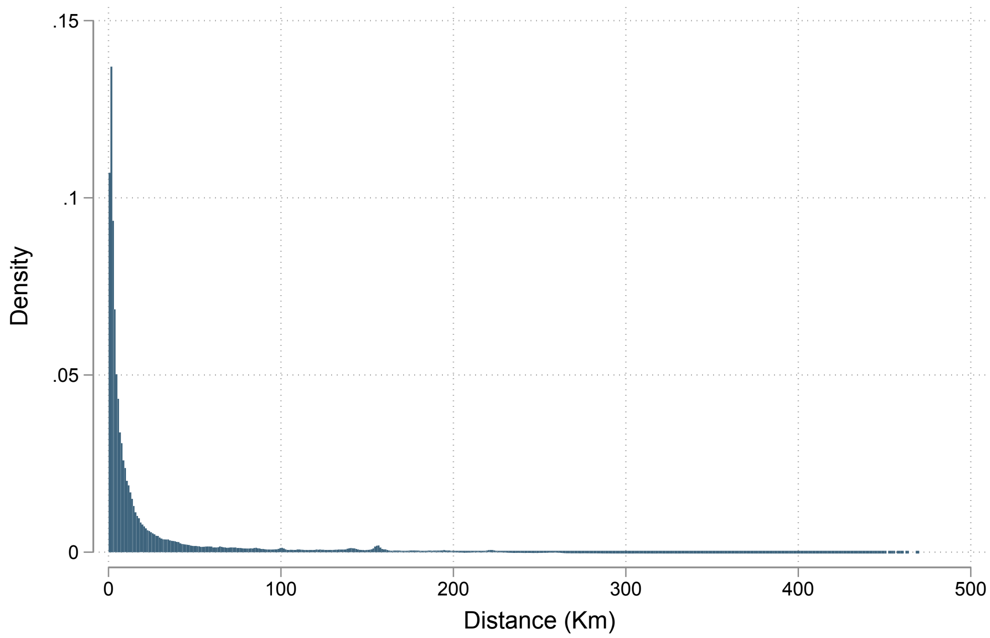


Figure A.5. Complete Histogram of Distances Between Origin and Destination Neighborhoods. In this figure, we plot the histogram of distances in kilometers between the origin and the destination neighborhoods for all Danish movers in our data.

A.5 Housing Prices

This section presents additional descriptives in order of appearance in Section 8.

Table A.1. Descriptive Statistics – Real Estate Transactions for Period 1993-2017

Variable	Mean	SD	Min	Max
Mean price (DKK M)	1.307	936	0.110	5.800
Space (sqm)	123.682	36.753	52.000	242.000
Number of rooms	4.299	1.249	1	8
Building age	53.886	37.745	0	1,016
Multifamily housing (share)	0.088	0.282	0	1

Note: Number of observations is 1,166,199.

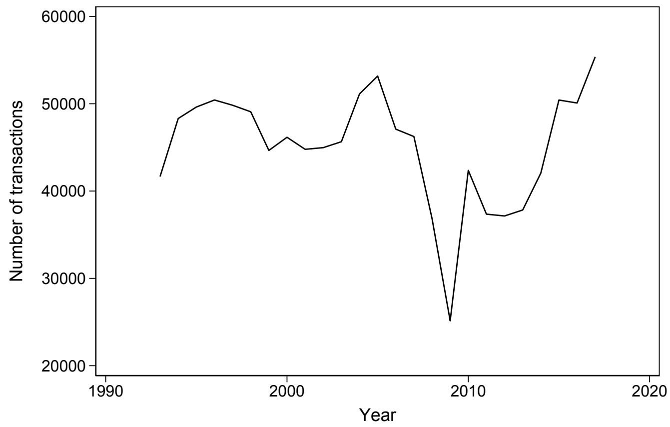


Figure A.6. Evolution of Number of Real Estate Transactions 1993-2017.

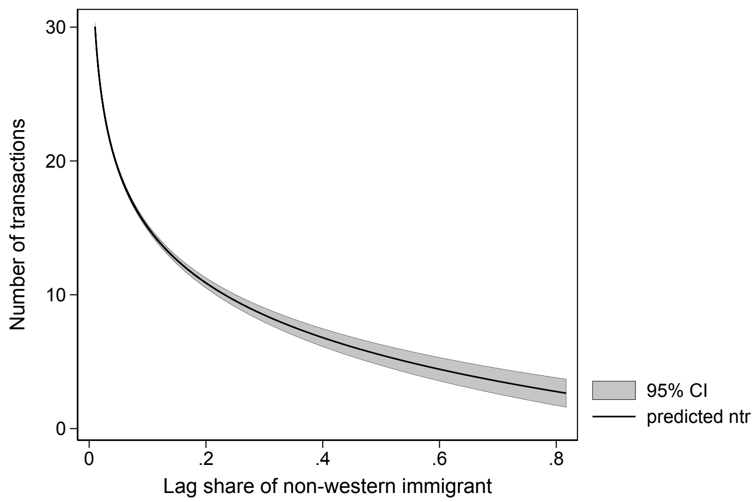


Figure A.7. Correlation Between the Share of non-Western Immigrants and the Number of Transactions at the Neighborhood Level.

A.6 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.[34] 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).

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

B Urban Amenities and Housing Price Index

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

B.1 Urban Amenities

We follow Glaeser et al. (2001) and create three types of amenities. First, we construct several consumer amenities such as food services, entertainment, and grocery. 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.[35] Amenities are determined by the total number of relevant establishments within a neighborhood. Figure B.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.

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.[36] Figure B.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 B.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

[35] 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.

[36] 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).

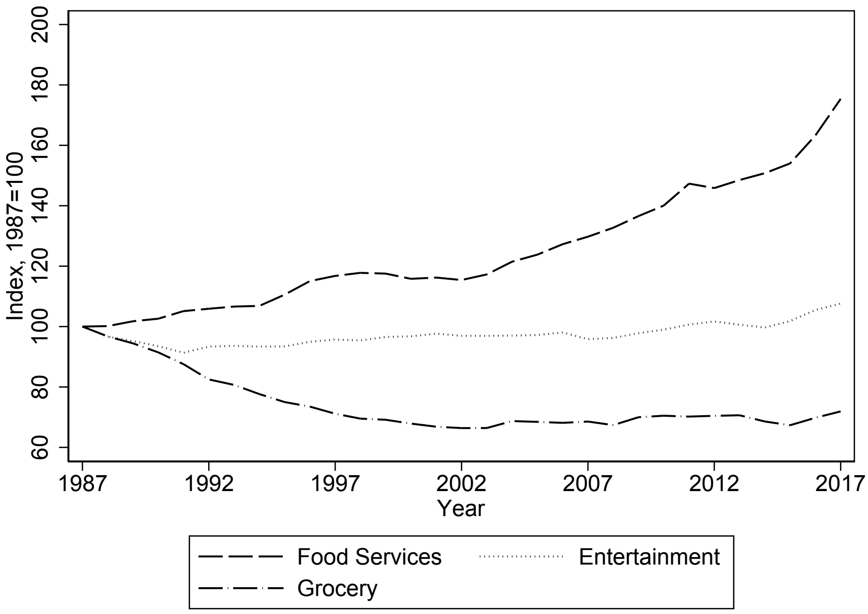


Figure B.1. Evolution of mean consumer amenities from 1987 to 2017. Consumer amenities are determined by the total number of relevant establishments within a neighborhood. The mean has been computed on a national basis.

of the amenities in the neighborhood is around one.

Table B.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.[37] Figure B.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 B.2. This is due to the centralization of

[37]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.

the banking sector. Table B.2 presents the summary statistics for amenities representing public goods, pooled across all years. As with Table B.1, there is significant variation in each amenity across neighborhoods. On average, there is less than one bank and primary school per neighborhood, but about one doctor per neighborhood.

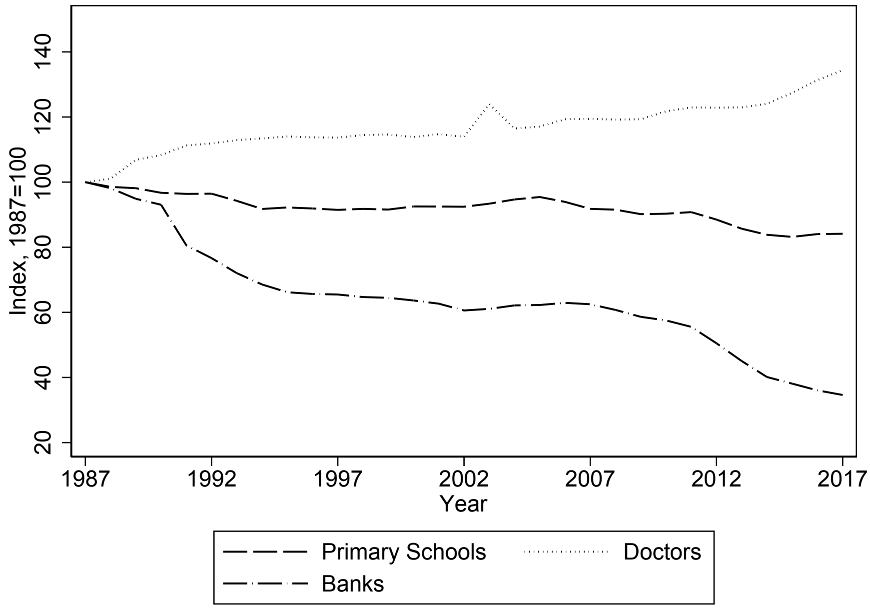


Figure B.2. Evolution of mean of selected public goods from 1987 to 2017. Consumer amenities are determined by the total number of relevant establishments within a neighborhood. The mean has been computed on a national basis.

Table B.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.[38] 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.

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

B.2 Housing Price Index

We estimate a house price index for neighborhoods for the period 1987-2017, using Danish administrative register data on real estate transactions. This data also includes structural dwelling attributes extracted from the Building and Dwelling Register (BBR), such as the size of the housing unit, the number of rooms, the type of unit (multifamily house, single-family house), and the age of the building. Since the real estate transaction data is sparse, we use an algorithm developed by Ahlfeldt et al. (2023) to predict a house price index for arbitrary spatial units from repeated cross-sections of geocoded microdata. We have information on the price paid for specific dwellings only for the period 1993-2017. For the period 1987-1992, we use the evolution of assessment of each dwelling and the estimates for the year 1993 to compute the housing price indices for each of the neighborhoods.

The mean sqm price is 12,464 DKK (std. dev. is 9,176). As depicted in Figure B.3, the sqm price initially decreased from 1987 to 1993, followed by an impressive 14-year streak of rising prices that eventually reached its peak in early 2007. This was followed by a sharp decline in prices until 2009, which soon stabilized. Since 2012, the sqm price has been on a strong upwards trend.

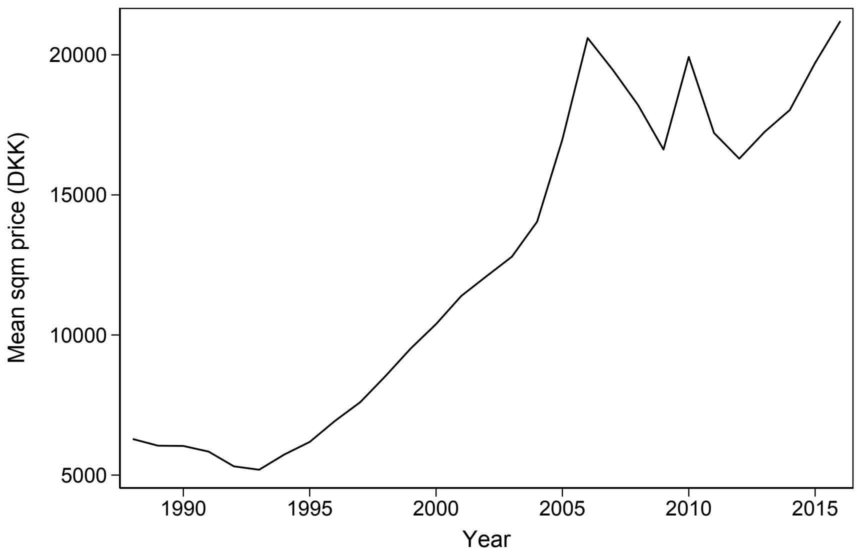


Figure B.3. Evolution of mean of the mean sqm price from 1987 to 2017.

Figure B.4 illustrates the housing price index across neighborhoods in Denmark in 2016. It is unsurprising that the highest housing prices are found in larger cities, particularly in the capital region and in the north of this region, which is seen as an attractive area. On the other hand, lower-priced homes are more widely available in western and southern Denmark.

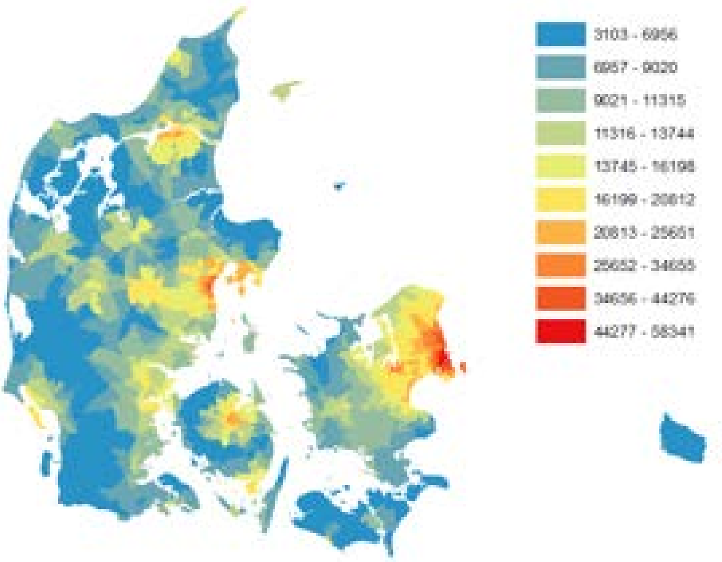


Figure B.4. Housing price index across neighborhoods in Denmark in 2016. This figure illustrates the price per square meter, measured in DKK/sqm, for neighborhoods across Denmark in 2016. 1 DKK ≈ 0.13 EUR.

C Additional Empirical Results

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

C.1 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 3, 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 C.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 neighborhoods.

C.2 Robustness to the Number of Units per Building

In Section 4, 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

Table C.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: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

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 C.2 and show that the coefficients are robust to the choice of the minimum number of apartments.

Table C.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: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

C.3 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 (1), 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 C.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 C.3. Robustness to the inclusion of additional trend variables

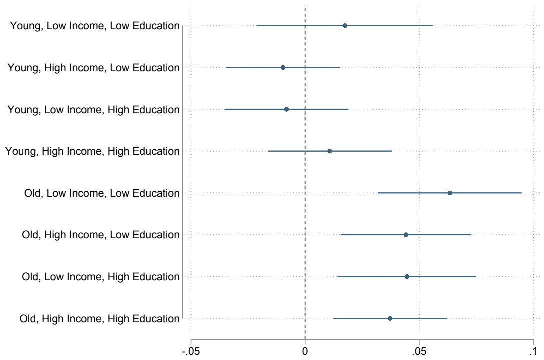
	Dependent Variable: Moved out of the Neighborhood in $t+1$			
	Full Sample	Single Family	Small Building or Complex (less than 10 units)	Large Building or Complex (10 or more units)
	(1)	(2)	(3)	(4)
Immigrant share (neighborhood)	0.017***	−0.001	0.003	0.032***
	(0.006)	(0.006)	(0.011)	(0.008)
N	53,332,175	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: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

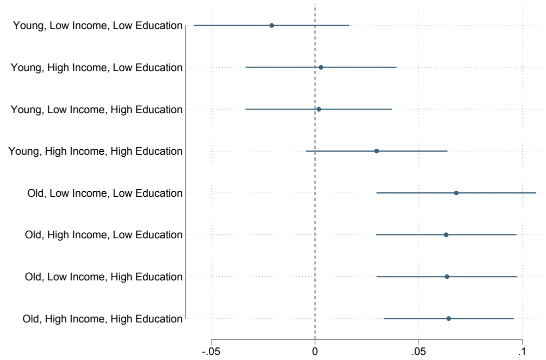
C.4 Robustness of heterogeneity results

When conducting our heterogeneity analysis in Section 4.5, 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 C.1, C.2 and C.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 4.5, but for two random partitions of the sample.

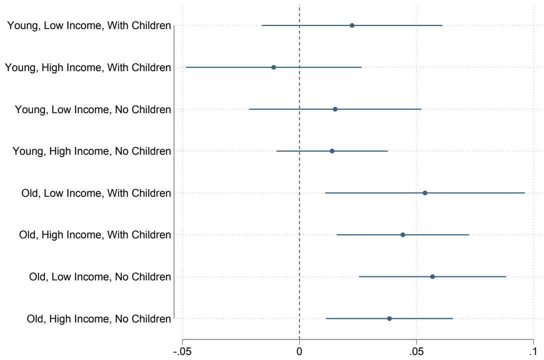


(a) Sample set 1

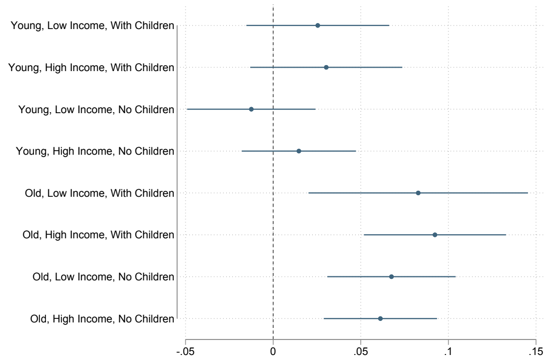


(b) Sample set 2

Figure C.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 4.5.1, but randomly splitting the sample into two parts.

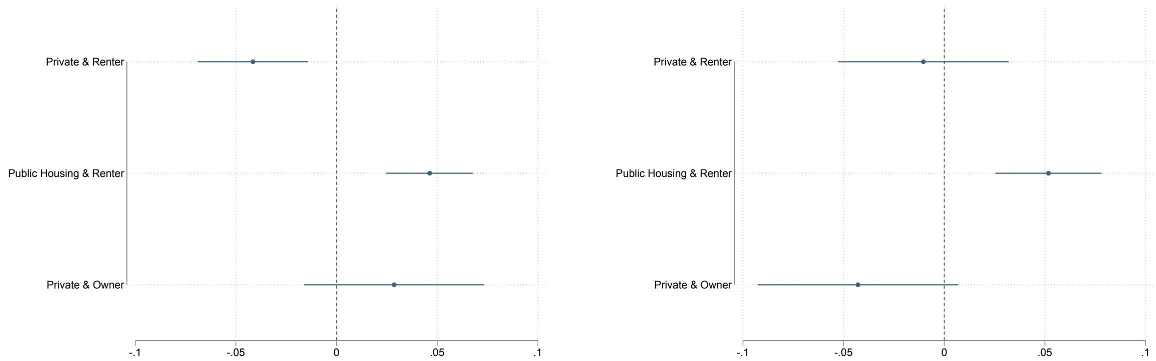


(a) Sample set 1



(b) Sample set 2

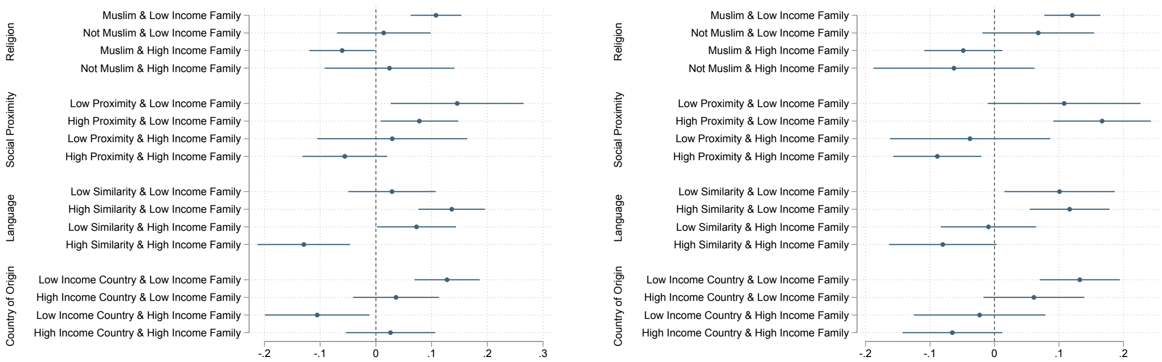
Figure C.2. Heterogeneity Analysis by Age, Income, and Children in the Household Splitting the Sample Randomly into Two parts. This figure presents the heterogeneity analysis as discussed in Section 4.5.2, but randomly splitting the sample into two parts.



(a) Sample set 1 (b) Sample set 2

Figure C.3. Heterogeneity Analysis by Homeownership Status Splitting the Sample Randomly into Two parts. This figure presents the heterogeneity analysis as discussed in Section 4.5.3, but randomly splitting the sample into two parts.

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

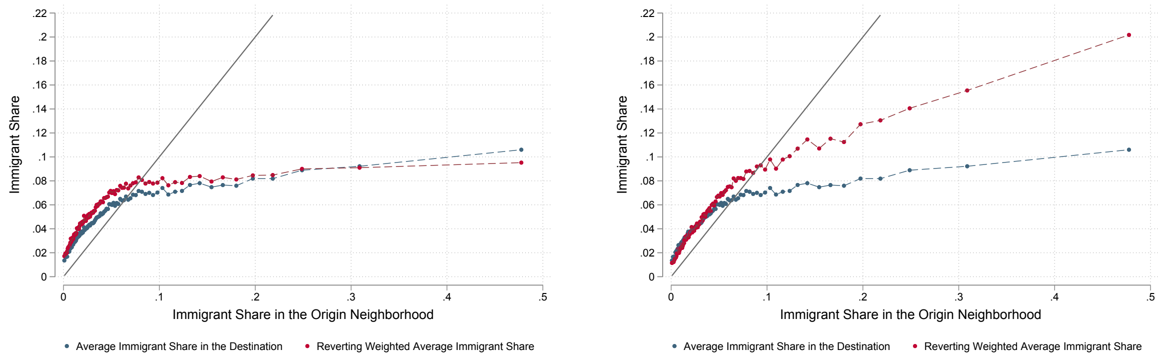


(a) Sample set 1 (b) Sample set 2

Figure C.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 5, but randomly splitting the sample into two parts.

C.5 Reverting Weighed Average Immigrant Share Under Different Spatial Decay Parameter Values

From Section 7, 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 C.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 C.5. Reverting Weighed Average Immigrant Share with Different Coefficients for the Spatial Decay Parameter, β.

D Instrumental Variable Construction

In this section, we discuss the details on the construction of the Arriver-Stayer-Markov Instrumental Variable (ASM-IV). We can write the number of *refugees* in neighborhood i in time t, $R_{i,t}$, as:

$$R_{i,t} = R_{i,t-1} - O_{i,t} + D_{i,t} + A_{i,t} \tag{D.1}$$

where the number of refugees in time t is the difference between the number of refugees from the previous year $R_{i,t-1}$, and the *outflow* of refugees that left the neighborhood i, $O_{i,t}$, plus the total arrival of refugees that were already in *Denmark* in the previous year but in any other neighborhoods $-i$ and choose to move to neighborhood i, $D_{i,t}$, and the total inflow of refugees that recently arrived in Denmark from *abroad* and went to neighborhood i, $A_{i,t}$.

Let us denote the difference between the previous year's number of refugees, $R_{i,t-1}$, and those that left neighborhood i, $O_{i,t}$, by $S_{i,t}$, as the total number of refugees that *stayed* in the neighborhood from the previous year.

$$R_{i,t} = S_{i,t} + D_{i,t} + A_{i,t} \tag{D.2}$$

In other words, the equation above states that, in any given year t, the total number of refugees in a neighborhood i is equal to the number of refugees that stayed there from the previous year, $S_{i,t}$, plus the new arrivals, from abroad $A_{i,t}$ or from other neighborhoods within Denmark $D_{i,t}$.

Now, considering the importance and distinctions in public versus private housing in the location of refugees in Denmark, we can further separate each of the components above into those that are living in Public housing, Pb, and those living in private housing Pv at period t.

$$\begin{aligned} R_{i,t} &= R_{i,t}^{Pb} + R_{i,t}^{Pv} \\ &= S_{i,t}^{Pb} + S_{i,t}^{Pv} + A_{i,t}^{Pb} + A_{i,t}^{Pv} + D_{i,t}^{Pb} + D_{i,t}^{Pv} \end{aligned} \tag{D.3}$$

In what follows, we discuss how exogenous each of these components is with respect to other factors that can be simultaneously associated with Danish mobility. We propose an approach that computes predicted inflows and outflows of refugees within Denmark based on national averages and pre-existing distribution of public housing and refugees.

Arrivals from abroad, $A_{i,t}$. Refugees who arrive in Denmark from abroad for the first time can be living in either public or private housing. As we discussed before, considering the Spatial dispersion policy in Denmark, it is fairly reasonable to assume that the arrival of refugees from abroad who are living in public housing is exogenous to the neighborhood characteristics that can influence both the allocation of refugees and the moving probability of Danish residents. Therefore, we assume that $A^{Pb}_{i,t}$, the new arrivals of refugees from abroad to public housing in the neighborhood i is exogenous.

However, the arrival of refugees from abroad to private housing may still be susceptible to endogeneity concerns, especially considering that some private landlords may refrain from renting out dwellings to refugees. Denote by $A^{Pv}_{i,t}$ the actual number of new refugee arrivals from abroad that live in private housing in neighborhood i. Denote by $\mathbb{P}(Z_i|Pv)$ the probability of any refugees moving to neighborhood i as their first place of residence, conditional on going to live in private housing. Then, we can write the number of arrivals in private housing as:

$$A^{Pv}_{i,t} = \mathbb{P}(Z_i|Pv)_t \cdot \sum_j A^{Pv}_{j,t} \tag{D.4}$$

To construct our instrument, we calculate a predicted number of new arrivals from abroad that eliminates potential confounding factors. To do this, we replace the actual probability $\mathbb{P}(Z_i|Pv)_t$ of refugees going to neighborhood i, by the share of refugees living in neighborhood i in 1987, $\hat{\theta}^R_{i,1987}$, and we exclude the actual arrivals to neighborhood i from the sum component. Therefore, the predicted number of new arrivals from abroad going to private housing becomes:

$$\hat{A}^{Pv}_{i,t} = \hat{\theta}^R_{i,1987} \cdot \sum_{j \neq i} A^{Pv}_{j,t} \tag{D.5}$$

Stayers, $S_{i,t}$. The number of refugees that stay from one year to the next may also be susceptible to endogeneity concerns, considering the potential attitudes of local natives toward the refugees.

Denote by $\mathbb{P}(M_{i,t})$ the probability that any of the refugees living in neighborhood i in $t-1$ chose to move to a different neighborhood in the next period t. So we could write the actual number of stayers in i as:

$$S_{i,t} = (1 - \mathbb{P}(M_{i,t})) \cdot R_{i,t-1} \tag{D.6}$$

The equation above states that the number of refugees that will stay in the neighborhood i is given by the probability of not moving away $(1 - \mathbb{P}(M_{i,t}))$, times the number of refugees in the neighborhood in the previous period. Now, the likelihood of refugees staying in a neighborhood may also be different depending on whether they live in public or private housing. Therefore, we assume that refugees have different likelihoods of moving to a different neighborhood depending on the type of housing they were living in the previous period (public or private). We can represent this by the equations below.

$$\begin{aligned} S_{i,t}^{Pb} &= (1 - \mathbb{P}(M_{i,t}|Pb_{t-1})) \cdot R_{i,t-1}^{Pb} \\ S_{i,t}^{Pv} &= (1 - \mathbb{P}(M_{i,t}|Pv_{t-1})) \cdot R_{i,t-1}^{Pv} \end{aligned} \tag{D.7}$$

Where we denote by $R_{i,t-1}^{Pb}$ and $R_{i,t-1}^{Pv}$ the total number of refugees located in neighborhood i, and living in public or private housing, respectively. $\mathbb{P}(M_{i,t}|Pb_{t-1})$ is the probability of a refugee moving to a different neighborhood between $t-1$ and t, conditional on living in public housing in $t-1$. $\mathbb{P}(M_{i,t}|Pv_{t-1})$ is the analogous probability of moving but for those refugees that were living in private housing.

One potential concern regarding our empirical strategy is whether the probabilities stated above can be endogenous to the mobility of natives. To address this potential endogeneity and construct our instrument, we compute the predicted number of refugees that stayed in public or private housing in neighborhood i using the following:

$$\begin{aligned} \hat{S}_{i,t}^{Pb} &= \left(1 - \hat{\rho}_{i,t}^{Pb}\right) \cdot R_{i,t-1}^{Pb} \\ \hat{S}_{i,t}^{Pv} &= \left(1 - \hat{\rho}_{i,t}^{Pv}\right) \cdot R_{i,t-1}^{Pv} \end{aligned} \tag{D.8}$$

where $\hat{S}_{i,t}^{Pb}$ is the predicted number of refugees that stayed in neighborhood i in public housing. $\hat{S}_{i,t}^{Pv}$ represents the stayers in private housing. $\hat{\rho}_{i,t}^{Pb}$ is the predicted measure for the probability of a refugee that was living in public housing moved to a different neighborhood between $t-1$ and t. We calculate this probability by the country-level frequency (excluding i) of refugees who moved to a different neighborhood in t and were living in public housing units in $t-1$. Analogously, $\hat{\rho}_{i,t}^{Pv}$ is the country-level (excluding i) frequency of refugees who moved to a different neighborhood in t and were living in private housing units in $t-1$, to proxy for the probability of refugees living in private housing changing neighborhoods.

Arrivals from other neighborhoods within Denmark, $D_{i,t}$. Refugees that were already in Denmark and decided to move to a particular neighborhood may be driven by factors that are confounding with the mobility of natives. To address this concern, we adopt the following strategy. From Equation (D.8), we can obtain the number of refugees that moved to a different neighborhood by multiplying the predicted probabilities of moving to the existing number of refugees for each type of housing.

$$\begin{aligned} N_{i,t,Pb} &= \mathbb{P}(M_{i,t}|Pb_{t-1}) \cdot R_{i,t-1}^{Pb} \\ N_{i,t,Pv} &= \mathbb{P}(M_{i,t}|Pv_{t-1}) \cdot R_{i,t-1}^{Pv} \end{aligned} \tag{D.9}$$

Where $N_{i,t,Pb}$ and $N_{i,t,Pb}$ are the number of refugees that moved to a different neighborhood, conditional on having lived in public or private housing, respectively.

Then, part of the moving refugees go to public housing, and part goes to private housing. Denote by $\mathbb{P}(Pb_t|Pb_{t-1}, M_t)$ as the probability of a refugee moving to public housing in t, conditional on living in public housing in $t-1$, having moved in the last period. Denote $\mathbb{P}(Pv_t|Pv_{t-1}, M_t)$ the analogous probability for refugees moving to private housing, conditional on having lived in private housing in $t-1$ and having moved. Therefore, we can obtain the expected number of refugees that moved to a different neighborhood and moved to either public or private housing as follows:

$$\begin{aligned} M_{i,t}^{Pb} &= \mathbb{P}(Pb_t|Pb_{t-1}, M_t) \cdot N_{i,t,Pb} + (1 - \mathbb{P}(Pv_t|Pv_{t-1}, M_t)) \cdot N_{i,t,Pv} \\ M_{i,t}^{Pv} &= \mathbb{P}(Pv_t|Pv_{t-1}, M_t) \cdot N_{i,t,Pb} + (1 - \mathbb{P}(Pb_t|Pb_{t-1}, M_t)) \cdot N_{i,t,Pb} \end{aligned} \tag{D.10}$$

Where $M_{i,t}^{Pb}$ and $M_{i,t}^{Pv}$ are the total number of refugees that moved out of neighborhood i, and went to live in public or private housing, respectively. From the total number of movers out of neighborhood i, The next step is to define the total number of refugees going to each other neighborhood. We begin by calculating the total number of refugees that moved to a different neighborhood in Denmark, except the ones that lived in the neighborhood i in the previous period. Denote by $\mathbb{P}(Z_i|Pb)_t$ the probability that any of the mover refugees choosing to move to neighborhood i as the *destination*, conditional on going to live in public housing. Also, denote by $\mathbb{P}(Z_i|Pv)_t$ the equivalent probability of refugees moving to neighborhood i, conditional on going to private housing. So we could write the expected number of new arrivals in i from within Denmark as follows:

$$\begin{aligned} D_{i,t}^{Pb} &= \mathbb{P}(Z_i|Pb)_t \cdot \sum_{j \neq i} M_{j,t}^{Pb} \\ D_{i,t}^{Pv} &= \mathbb{P}(Z_i|Pv)_t \cdot \sum_{j \neq i} M_{j,t}^{Pv} \end{aligned} \tag{D.11}$$

Where $D_{i,t}^{Pb}$ is the number of refugees choosing neighborhood i as a destination, conditional on going to live in a public house in t. Analogously, $D_{i,t}^{Pv}$ is the number of refugees choosing neighborhood i as a destination, conditional on going to live in private housing in t. In constructing our instrumental variable, another concern may arise regarding the probabilities that govern the destination choices of domestic movers. To address this issue, we replace these probabilities with the pre-existing shares of public housing or refugees located in a neighborhood in 1987. We can write the predicted number of

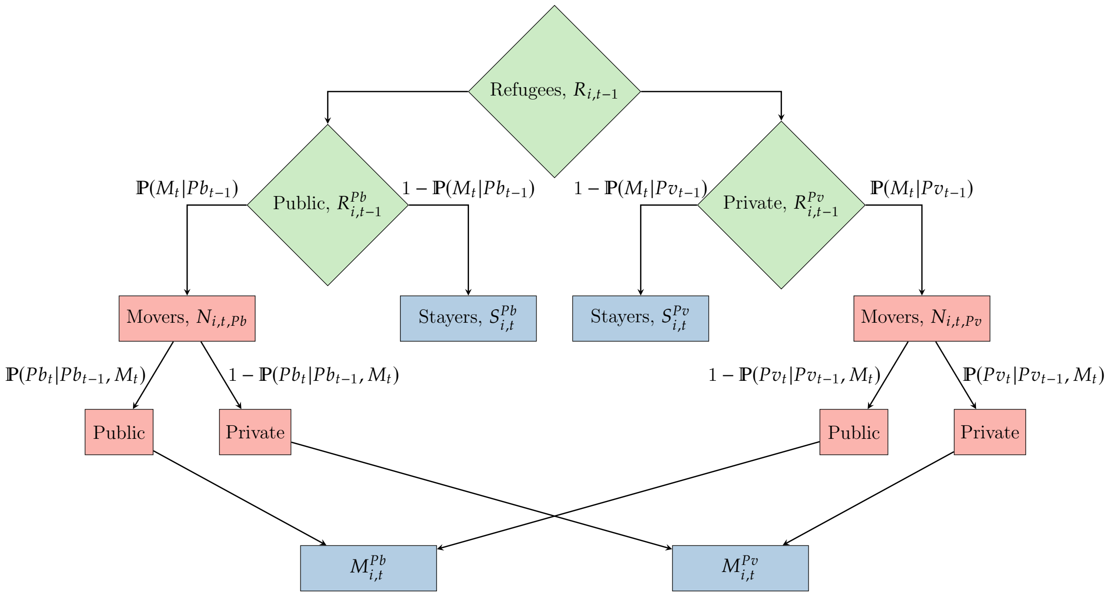


Figure D.1. IV Design - Structure in computing the people moving within Denmark

arrivals from within Denmarks to neighborhood i as follows:

$$\begin{aligned} \hat{D}^{Pb}_{i,t} &= \hat{\theta}^{Pb}_{i,1987} \cdot \sum_{j\neq i} \hat{M}^{Pb}_{j,t} \\ \hat{D}^{Pv}_{i,t} &= \hat{\theta}^{R}_{i,1987} \cdot \sum_{j\neq i} \hat{M}^{Pv}_{j,t} \end{aligned} \tag{D.12}$$

where $\theta^{Pb}_{i,1987}$ is the share of Denmark's public housing units located in neighborhood i in the reference year, 1987. As we discussed before, $\theta^{R}_{i,1987}$ is the share of refugees in Denmark that were living in the neighborhood i in the reference year. The intuition is that refugees moving to public housing will be allocated to neighborhoods in Denmark according to the pre-existing distribution of public housing units, while refugees moving to private housing will be allocated to neighborhoods according to the pre-existing distribution of refugees in Denmark.

Figures D.1 and D.2 illustrate the structure of the instrument design and how each component is used to calculate the predicted number of refugees in a neighborhood.

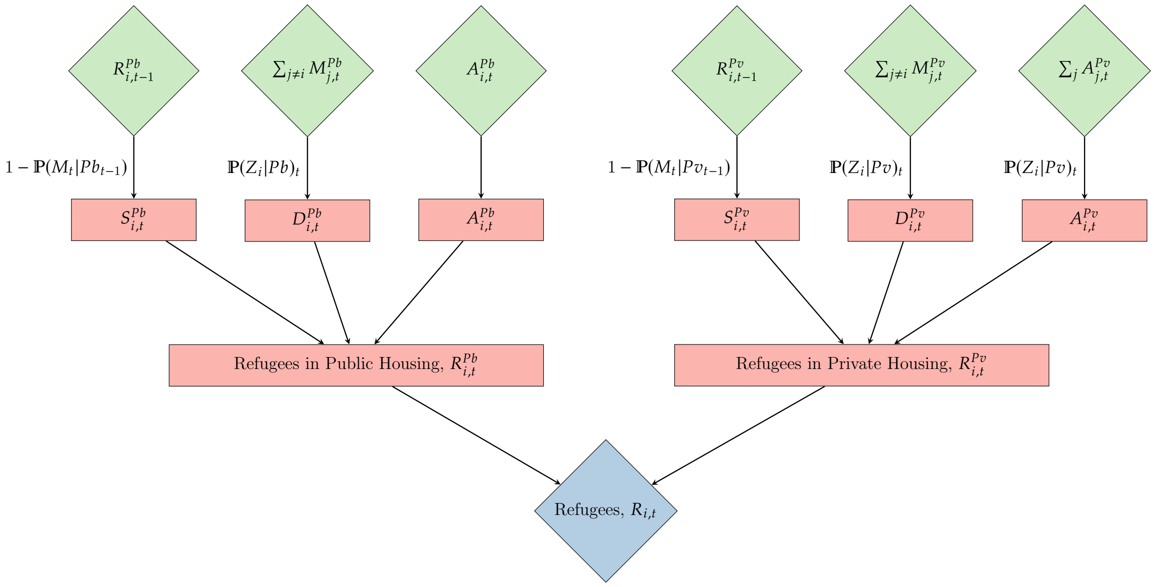


Figure D.2. IV Design - Structure for Computing Refugee Presence