

## 1 Introduction

Immigrants who move to a new country during early childhood tend to do better in adulthood along a host of dimensions: they achieve better educational and labor market outcomes (Böhlmark 2008; Hermansen 2017; Alexander and Ward 2018; Lemmermann and Riphahn 2018; Ansala et al. 2019), they enjoy better health (Van den Berg et al. 2014), and they exhibit higher levels of social and political integration (Åslund et al. 2015; Andersson et al. 2025). Immigrants, more generally, also tend to live in segregated neighborhoods and this holds even decades after arrival and across host countries (e.g. Cutler et al. 2008; Malmberg et al. 2018). In this paper, we ask whether age at arrival affects residential segregation in adulthood. While previous studies focus on non-residential outcomes or use the minority share of a neighborhood as a measure of residential integration, the more complex outcome of residential segregation has largely been ignored.

We use high-quality register data from Sweden and apply a siblings design to immigrant cohorts born between 1974 and 1987 to study whether a younger age at arrival has an impact on the level of segregation in the neighborhoods they reside in at age thirty and how that effect differs for refugees and non-refugees. In addition, we explore potential integration channels through which earlier age at arrival may lead to lower segregation outcomes.

We use the neighborhood contribution to the dissimilarity index as a novel dependent variable (DV). This widely used index of urban segregation is the sum of each neighborhood's absolute divergence from the municipality-level immigrant average share. The higher the divergence, the more the neighborhood contributes to total urban segregation. Using each immigrant's neighborhood contribution as DV, we exploit variation in age at arrival between siblings to estimate the effect of arriving at different ages during childhood (and before sixteen) relative to a reference group that arrives between ages 0 and 3. The within-family analysis enables us to address potential selection bias stemming from the fact that parents with better unobserved characteristics may move abroad when their children are younger.<sup>1</sup>

Our overall finding is that compared to immigrant children arriving between the ages of 0 and 3, immigrant children arriving later live in more segregated neighborhoods. This result is even stronger for refugee than for non-refugee immigrants. Moreover, the timing of effects looks different for the two groups: for refugees, arriving at age 4 or later increases the neighborhood contribution, while for non-refugees, the positive effect starts emerging only around age 11. At that same age, the effects increase further for refugees. This suggests the existence of a critical age for non-refugees, and multiple sensitive periods for refugees. Previous research finds that for outcomes such as education and language attainment, the critical age is 8-10 (Böhlmark 2008; Basu 2018) or even 6 (Lemmermann and Riphahn 2018).

Spatial assimilation theories suggest that residential integration is the end-product of the integration process (Massey and Denton 1988). We therefore expect to see similar patterns of

---

<sup>1</sup> We note, however, that such issues are likely to be less prevalent for refugees, who are more likely to move so as to escape violence and conflict and thus have less control over the timing of their moves.

age at arrival effects on other integration outcomes. In particular, we would expect age at arrival effects to be flat until age 11 for non-refugees, and increasing in age for refugees, with stronger effects after age 11. To test for this, we apply our analysis to the following outcomes: income rank, education and intermarriage. We find that the age at arrival effects on these other outcomes are more similar in magnitudes for the two groups, despite the differences in segregation effects. At the same time, the effects level off for non-refugees around ages 10-11, but not for refugees. This result suggests that the mechanisms underlying segregation operate differently for refugees versus non-refugees.

To probe this hypothesis more thoroughly, we conduct a decomposition analysis in the style of Heckman et al. (2013) to quantify how much of the effect of age at arrival on neighborhood integration goes through these three important mechanisms. For non-refugees, residential segregation largely reflects economic integration—consistent with spatial assimilation theories linking income and education to residential mobility. Among refugees, both economic and social integration matter, yet substantial unexplained gaps persist despite similar labor market and intermarriage patterns. This could be related to refugees facing more structural barriers in accessing good jobs (Helgesson et al. 2019) or good housing (Andersson et al. 2010) when compared to non-refugees. However, we interpret this analysis as descriptive and suggestive rather than causal; identifying the precise mechanisms requires future research.

Our paper makes two main contributions. First, we introduce residential segregation as an outcome in the age-at-arrival literature. While previous studies have shown that earlier arrival leads to better outcomes across multiple dimensions, residential segregation has remained largely unexplored. Integration is a multidimensional process, and understanding how it unfolds along these multiple dimensions is important for developing adequate policy responses (see Harder et al. 2018, Aksoy et al. 2023). Second, we focus specifically on refugees rather than economic migrants. Brell et al. (2020) argue that this distinction matters because refugees face fundamentally different circumstances. Unlike children of economic migrants whose parents select the destination based on economic opportunities, refugee children arrive in host countries their parents did not necessarily choose. Refugee children often also experience disrupted schooling due to conflict, displacement, or time spent in refugee camps. They are more likely to have been exposed to violence, persecution, and traumatic experiences during displacement, potentially affecting their mental health and educational outcomes. The importance of these factors varies with age at arrival: younger children may face fewer difficulties navigating these challenges compared to older arrivals. Until recently, most datasets did not distinguish between refugees and other immigrants (Brell et al. 2020). Swedish administrative data reliably records refugee status, addressing a key data limitation in the existing literature.

The study most closely related to ours is Åslund et al. (2015), but our work differs in two key ways. While we focus on recent cohorts of refugees, Åslund et al. (2015) examine the children of earlier cohorts of labor immigrants, primarily from other Nordic or other European countries. Additionally, Åslund et al. (2015) interpret their findings as suggesting that economic factors play a marginal role in shaping segregation later in life, with cultural identity being more

influential. In contrast, our analysis shows that for refugees, cultural identity (proxied by intermarriage) and economic factors (education and income) contribute equally to explaining segregation outcomes. For non-refugees, economic factors play a more dominant role.

This research note first introduces the unique data which allow implementing our empirical strategy, before presenting the results in three steps: effect on residential segregation, effects on other integration outcomes and a decomposition of residential effects into labor market and social integration channels.

## 2 Data, empirical strategy and descriptive statistics

### 2.1 Data and sample selection

We use Swedish geo-coded register data from the GeoSweden database, which contains information on all residents in Sweden. The data is collected on a yearly basis from 1990 to 2017 and consists of variables from the population and tax registers. Importantly for our study, it also contains information on the country of birth, reason for and year of immigration. It additionally includes detailed geographic information on residential location.

Our sample consists of immigrant children born between 1974 and 1987 and whose age upon arrival in Sweden is between zero and fifteen.<sup>2</sup> We measure outcomes at age 30, similarly to prior studies (e.g. Hermansen 2017), an age by which most individuals have made at least one independent residential choice and are likely to be relatively settled. Measuring outcomes at older ages would substantially reduce the sample size and limit the number of immigrant cohorts that we can include. About 87% of individuals in our sample live outside the parental home at age 30, indicating that residential patterns at this age primarily reflect adult location decisions. Because outcomes are observed at age 30, the analysis necessarily focuses on immigrants who remain in the host country up to that age, excluding those who return to their country of origin earlier.

We classify immigrant children into three categories: all immigrants, refugees and non-refugees. An individual is considered a refugee if either their own permit is a refugee permit or, absent this information, if they have at least one parent classified as a refugee. A non-refugee is an individual who does not fulfill these criteria.<sup>3</sup> The "all immigrants" category pools together

---

<sup>2</sup> The earliest cohort that we can observe at age 16 is born in 1974, whereas the youngest cohort we can observe at age 30 is born in 1987. Hence, these data restrictions inform our choice of the cohorts under study. The age at arrival variable comes primarily from the in-migration register, which is available from 1990 to 2017. For those arriving before 1990, we use a variable from the income register (Louise) that gives the latest year of immigration. We take the value of this variable when the child first enters the Louise register, at age 16.

<sup>3</sup> We have permit information for at least one of either the child, the mother or the father for 89.71% of the sample. For the remaining 5,041 observations (or 10.29% of the sample) the permit information is missing. Note, however, that in our data, missing permit information may simply indicate that the individual does not require a permit. We categorize observations with missing permit data as non-refugees. Our results are robust to reclassifying them as refugees instead.

refugees and non-refugees. Regardless of refugee status, all immigrants are born abroad to foreign-born parents.

## 2.2 Outcomes

### *Residential segregation*

We are interested in the degree of residential segregation in the neighborhood, where an immigrant who arrived in Sweden as a child resides at age 30. Our neighborhood measure is the so-called DeSO (*Demographic Statistical Area* or *demografiska statistikområden* in Swedish), an administrative unit defined by Statistics Sweden such that the boundaries follow, to the extent possible, streets, waterways and railways. There are approximately 6,000 DeSOs in Sweden, with population ranging from 700 to 2,700 and thus slightly smaller than US Census Tracts, for example.<sup>4</sup> DeSOs are often used in Swedish migrant segregation research when the goal is to capture lived experiences of segregation at smaller geographical scales (Cederström et al. 2025). To measure segregation, we use the well-established dissimilarity index (Duncan and Duncan 1955), which captures the evenness dimension of segregation (Massey and Denton 1988). For robustness, we also report results using the isolation index, which reflects the exposure dimension. The dissimilarity index is generally defined as:

$$D = \frac{1}{2} \sum_{i=1}^N \left| \frac{a_i}{A} - \frac{b_i}{B} \right| \quad (1)$$

where  $D$  is the Dissimilarity Index,  $N$  is the number of neighborhoods in a municipality,  $a_i$  represents the number of immigrants in the  $i$ -th neighborhood,  $A$  is the total number of immigrants in the municipality,  $b_i$  represents the number of natives in the  $i$ -th neighborhood, and  $B$  is the total number of natives in the municipality. The index ranges between 0 and 1 and can be interpreted as the proportion of people in a group who would have to move in order for each neighborhood to have the same proportion of that group as the municipality as a whole. It hence measures the unevenness of the immigrant distribution across a municipality in a given year (Massey and Denton 1988).

We use the neighborhood contribution to the dissimilarity index as our outcome variable. For each neighborhood  $i$ , this corresponds to  $|a_i/A - b_i/B|$  in equation (1) above. This measure captures how much each neighborhood contributes to overall municipal segregation. Individuals who live in the same neighborhood  $i$  have identical values of this contribution measure. If such a neighborhood were a perfect copy of the municipality, the contribution would be zero. Higher values, i.e. diverging from the municipality average, indicate residence in neighborhoods that contribute more to municipal segregation. To compute  $D$  and the neighborhood-level contribution, we restrict the groups ( $a_i$ ,  $b_i$ ,  $A$  and  $B$ ) to be between 18 and 60 and define the group of immigrants as born abroad to foreign-born parents.

---

<sup>4</sup>US Census Tracts have a population size between 1,200 and 8,000 people (see [https://www.census.gov/programs-surveys/geography/about/glossary.html#par\\_textimage\\_13](https://www.census.gov/programs-surveys/geography/about/glossary.html#par_textimage_13), accessed on Oct 23, 2025, for more details).

In our pooled sample, the municipality-level  $D$  ranges from 0.009 to 0.99, from almost zero to complete segregation, with a median value of 0.74, while neighborhood-level contribution ranges from 0 to 0.36, with a median value of 0.007 and a mean of 0.02 (cf. Table 1, Panel A), suggesting fairly uneven contributions across neighborhoods. This difference in the amount of variation is important for interpreting the magnitude of our results below.

Using this neighborhood-level measure is preferable to using the aggregate municipal dissimilarity index  $D$  for several reasons. First, our identification strategy requires within-family variation in outcomes (see section 2.3). Since  $D$  is calculated at the municipal level, siblings living in the same municipality would have identical values, eliminating the variation needed for identification. Our measure varies across neighborhoods within municipalities, capturing differences in where siblings reside. Moreover, segregation research highlights stark intra-urban differences, which the Swedish data quality uniquely allows to capture. Second, our measure reflects where individuals actually reside within their municipality, not just the overall segregation level of their municipality. Two individuals in municipalities with similar  $D$  values may live in very different neighborhood contexts—one in a highly segregated enclave, another in a mixed area. Our measure captures these residential sorting patterns. Finally, as segregation measures are much discussed, we do also show results for the simpler neighborhood share of immigrants and the alternative isolation index, proxying the exposure dimension of segregation, for robustness.

### *Other outcomes*

In a complementary analysis, we quantify the extent to which the residential segregation outcomes work through labor market and social integration. Therefore, we also study the effect of age at arrival on income rank, years of education, marriage and intermarriage. An individual's income rank is the percentile rank based on his or her position in the national distribution of incomes relative to all individuals in the same birth cohort. The income definition includes labor income and income from self-employment. The years of education variable is constructed by translating educational levels into corresponding years of education. Marriage is defined as either married or cohabiting with children. We consider an individual to be intermarried if their partner is born in Sweden.

## **2.3 Empirical strategy**

We use the samples of immigrant children as defined in section 2.1 to estimate the following equation:

$$y_{ij} = \alpha + \sum_{a=4}^{15} \beta_a I(a_{ij} = a) + \mu_{\text{first-born}} + \theta_{\text{female}} + \phi_j + \eta_{ij}$$

where  $y_{ij}$  is the outcome of child  $i$  in family  $j$ ,  $a_{ij}$  is the child's age at arrival in Sweden,  $\phi_j$  is a family fixed effect that captures unobserved family characteristics that are common to all siblings in the same family and constant over time, and  $\eta_{ij}$  is the error term.<sup>5</sup> Those that arrive at ages 0-3 constitute the reference group.

Our empirical strategy addresses the concern that parents with better unobserved characteristics (in terms of, for example, motivation, parenting skills, and other variables that might be correlated with the outcome variables but that are not observed in the data) may migrate to a larger extent when their children are young. Identification of the  $\beta_a$  coefficients of interest comes from variation in age at arrival between siblings. Using this approach, the coefficients reflect the combined effect of age at arrival and length of stay in Sweden.<sup>6</sup> We follow the previous literature that highlights the importance of birth order effects and add a dummy for first-born children (Böhlmark, 2008). The *female* dummy captures gender differences in the outcomes we consider. Table 1 shows summary characteristics for each immigrant group in the siblings sample.<sup>7</sup> Focusing on Panels B and C, we see that, on average, refugees and non-refugees live in neighborhoods that are similar in terms of segregation. In terms of labor market integration, refugees have on average a higher income rank and more years of education. However, they are more likely to be married and less likely to be married to a native partner. On average, refugees arrive when they are 1.7 years older than non-refugees.<sup>8</sup>

---

<sup>5</sup> We identify siblings through their mother when maternal identifiers are present in the registers; when maternal identifiers are unavailable, we identify siblings through their father.

<sup>6</sup> There is very little variation in years of arrival between siblings, hence results are very similar when using a specification that adds year of arrival dummies.

<sup>7</sup> Table A.1 shows the analogous summary statistics for the full sample; there are no major differences between the siblings samples and the full samples, in either of the groups we study.

<sup>8</sup> Table A.2 shows the breakdown of country of origin, by refugee status. Generally speaking, non-refugees are primarily from other Nordic and European countries; whereas refugees tend to be from countries going through conflicts during the sample period.

Table 1: Summary statistics for the siblings sample

	Mean	Std. dev.	No. of obs.
<i>Panel A: All immigrants</i>			
Neighborhood contribution to dissimilarity index	0.020	0.032	48,980
Income rank	45.551	30.716	48,980
Years of education	12.297	2.269	48,564
Married	0.425	0.494	48,980
Intermarried	0.312	0.463	20,816
Female	0.473	0.499	48,980
First-born	0.377	0.485	48,980
Age at arrival	8.722	3.783	48,980
<i>Panel B: Refugees</i>			
Neighborhood contribution to dissimilarity index	0.020	0.033	38,422
Income rank	46.061	30.741	38,422
Years of education	12.342	2.276	38,111
Married	0.433	0.496	38,422
Intermarried	0.263	0.440	16,644
Female	0.472	0.499	38,422
First-born	0.362	0.481	38,422
Age at arrival	9.094	3.595	38,422
<i>Panel C: Non-refugees</i>			
Neighborhood contribution to dissimilarity index	0.017	0.028	10,558
Income rank	43.695	30.555	10,558
Years of education	12.135	2.237	10,453
Married	0.395	0.489	10,558
Intermarried	0.508	0.500	4,172
Female	0.477	0.500	10,558
First-born	0.432	0.495	10,558
Age at arrival	7.366	4.126	10,558

*Notes:* This table reports summary statistics for all immigrants, refugees and non-refugees in the siblings sample, respectively. Children are born between 1974 and 1987. We classify a child as a refugee if either their own permit is a refugee permit or, absent that information, if they have at least one parent classified as a refugee. The dissimilarity index is the absolute value of the individual component for each  $i$ -th neighborhood in equation 1. Intermarriage is marriage to a Swedish-born partner.

*Source:* Own calculations on data from the GeoSweden database.

### 3 Results

We present our results in the following three sections. In section 3.1, we first show the effects of age at arrival on residential segregation at age 30, defined above as the neighborhood-level contribution to the municipality-level dissimilarity index. In order to examine the extent to which the effects on residential segregation work through labor market and social integration, we then estimate the effects of age at arrival on income rank, educational attainment and

marriage and intermarriage in section 3.2 separately. Finally, we decompose the main effect estimated in section 3.1 into parts attributable to the different channels in section 3.3.

### 3.1 Effects on residential segregation

Figure 1 plots the  $\beta_a$  coefficients obtained when estimating equation (2) with the neighborhood contribution to the dissimilarity index as the dependent variable. Overall, we see that immigrants who arrive later live in more segregated areas at age 30. Relative to those arriving at age 0-3 (our reference category), whose neighborhood contribution to dissimilarity is 0.014 (Table A.3), immigrants arriving at age 15 live in neighborhoods that contribute an additional 0.009 to municipal segregation, which represents 64% of the baseline. Another way to interpret this magnitude is to note that it represents 0.28 standard deviations in the immigrant distribution of neighborhood contributions (Table 1, Panel A), indicating a meaningful difference in residential outcomes. While individual neighborhood contributions are by construction small in absolute terms, these effects aggregate across the many neighborhoods within municipalities where immigrants concentrate.<sup>9</sup>

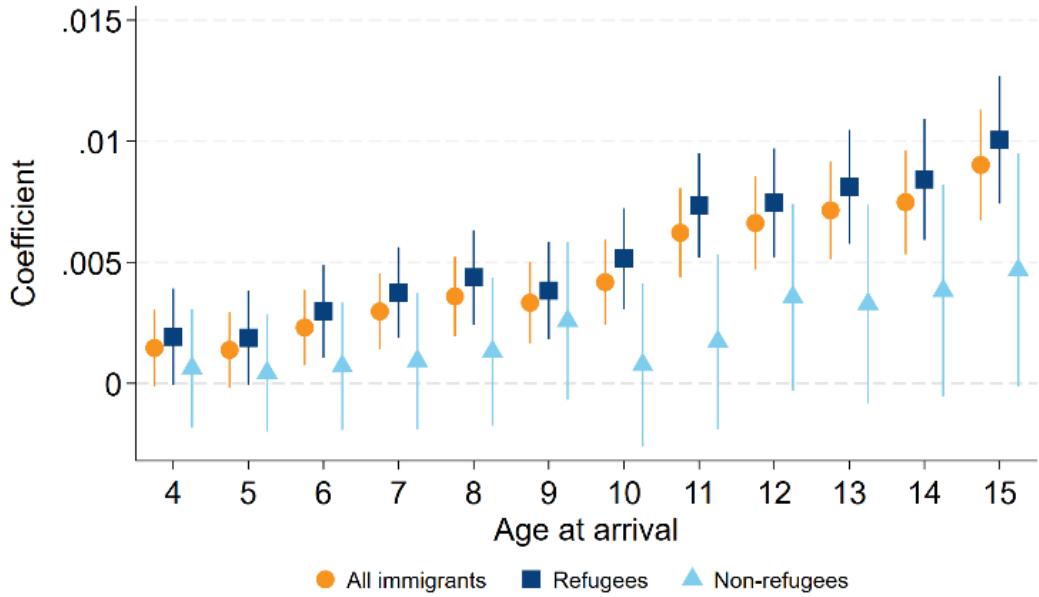
Both refugees and non-refugees show a marked change in slope at age 11, with noticeably different patterns before this threshold.<sup>10</sup> For refugees (dark blue squares), non-zero effects start immediately at age 4 and increase roughly linearly until age 11, after which the slope steepens. This suggests that each year of delayed arrival matters for refugees, with effects intensifying after age 11. In contrast, for non-refugees (light blue triangles), effects remain largely flat until age 11, when they start increasing slightly. This flat initial pattern indicates that for non-refugee individuals, barriers to residential integration emerge only for those arriving after age 11. This result is even more striking when we look at Table A.3, which shows that both refugees and non-refugees arriving at ages 0-3 live in neighborhoods that contribute similarly to the municipality-level segregation. The effect on all immigrants (orange dots) is therefore primarily driven by the effect on refugees, with non-refugee effects remaining roughly half the size of refugee effects even at age 15.

---

<sup>9</sup> These results may be downward biased in the presence of spillovers across siblings: since those who arrive at older ages settle in more segregated areas, if younger siblings have a preference for living close to their older siblings, they may choose a more segregated area than they otherwise would in the absence of these spillovers.

<sup>10</sup> We note, however, that confidence intervals overlap across all ages of arrival.

Figure 1: Effect of age at arrival on the neighborhood contribution to dissimilarity index



*Note:* The figure shows the  $\beta_a$  coefficients obtained when estimating equation (2) and their corresponding 95% confidence intervals.

*Source:* Own calculations on data from the GeoSweden database.

### 3.1.1 Robustness checks

In this subsection, we examine the robustness of our main results to alternative dependent variables and alternative definitions of our samples.

#### *Alternative dependent variables*

Our dependent variable captures how much a neighborhood contributes to overall municipal uneven dimension of segregation. We now analyze to what extent our results are sensitive to using two alternative dependent variables: i) the neighborhood contribution to the isolation index<sup>11</sup>, which captures the exposure or interaction dimension (Figure A.1) and ii) the share of immigrants in a neighborhood, which simply captures the composition of a neighborhood (Figure A.1).

In Figure A.1, we see that for non-refugees, age at arrival does not matter for the neighborhood contribution to the isolation index: the coefficients are 0 across all ages, which could be

---

<sup>11</sup> The formula for the neighborhood contribution to the isolation index is:  $(a_i/A)(a_i/t_i)$ , where, as before,  $a_i$  represents the number of immigrants in the  $i$ -th neighborhood,  $A$  is the total number of immigrants in the municipality and  $t_i$  is the total population in the neighborhood.

explicable by their lower segregation rates and larger group size on which the isolation index depends in contrast to the dissimilarity index. For refugees, by contrast, we see a flat pattern and coefficients close to 0 up to the age of 11 – the same threshold as before, when the coefficients increase slightly. For this measure of segregation, the magnitudes are lower than in our previous results. For example, arriving at age 15 increases the neighborhood contribution by 0.004, which is 40% of the baseline mean of 0.01, compared to 64% of the baseline mean in the neighborhood contribution to dissimilarity. As the total sum of the isolation index also depends on the size of the minority group, however, the contribution numbers are not directly comparable.

Figure A.2 reveals a mostly flat and negative pattern across all groups. In other words, later arrivals do not systematically sort into neighborhoods with different compositions relative to earlier arrivals. Across the board, from age 5 onwards, immigrants reside in neighborhoods with *lower* immigrant shares than those arriving at ages 0-3. Together with the results from section 3.1, our results show that later-arriving immigrants live in neighborhoods with higher contributions to the dissimilarity index but slightly lower neighborhood immigrant shares.

To make sense of these results, we note the following statistics: those that arrive at ages 0-3 live in neighborhoods with on average 34% immigrants (Table A.3), but in municipalities with on average 27% immigrants. Those that arrive at age 15, for example, live in neighborhoods with 38% immigrants, in municipalities with on average 23% immigrants. Later arrivals tend to settle in immigrant-heavier neighborhoods within these lower-immigrant municipalities, producing high neighborhood-level dissimilarity. In contrast, early arrivals are more likely to reside in municipalities with higher overall immigrant shares, where neighborhood compositions are closer to the municipal average and thus contribute less to overall segregation. One plausible explanation is that later arrivals are more likely to stay in their initial location; if those locations were assigned through refugee dispersal policies, for example, we would expect the municipality immigrant average to be smaller. Using the neighborhood contribution to the dissimilarity index as DV factors in the divergence from the municipality mean and thus the within-municipality differences across neighborhoods, which the simple share foreigners does not capture.

#### *Alternative sample definitions*

Our identification strategy relies on variation in age at arrival across siblings. However, for siblings with large age gaps, the older child may experience the parents' early, less stable integration years, while the younger child may grow up in a more established environment. Because such differences are time-varying and correlated with age at arrival, they can introduce attenuation bias.

We therefore assess whether our results change when we exclude sibling pairs with an age gap larger than 5 years.<sup>12</sup> Figures A.3a (refugees) and A.3b (non-refugees) show that the coefficients increase for both groups (with no changes in precision), but they increase relatively more for non-refugees, suggesting the presence of attenuation bias in our baseline specification. Nonetheless, the result that the effect of age at arrival on residential segregation is stronger for refugees than it is for non-refugees stands.

### **3.2 Effects on labor market, educational, and social integration**

The earlier immigrant children arrive in a new country, the more time they have to build country-specific knowledge (e.g. different types of networks, language, cultural habits, institutional knowledge). This country-specific knowledge might also affect other forms of (integration) outcomes that, in turn, might affect residential integration. Here we examine the effects on three other important margins: labor market, educational, and social integration.

Across all outcomes, we see very strong age at arrival effects for both refugees and non-refugees, but with slightly differing patterns. First, when it comes to income rank (panel a), refugees experience steadily increasing negative effects of age at arrival, with a drop of up to 15 percentile ranks lower in the national income distribution for those who arrive at age 15 compared to those who arrive at ages 0-3. For non-refugees, the coefficients are very similar in magnitude but only up to age 10, when they level off.<sup>13</sup>

Panel (b) shows a rather flat pattern for both groups up to the ages of 7-8, when the coefficients start to noticeably drop. For refugees, they continue to drop rather linearly up to the age of 15, when the effect is 0.8 years of education lower than the reference category. For non-refugees, effects are constant from age 8 to 11, when they drop again. From age 10 onward, the effects are always less negative than for refugees. These results echo previous findings in the literature identifying critical ages around the time students enter school (e.g. Böhlmark 2008, among others).

Panel (c) reveals that the probability of being married at age 30 increases with age at arrival. Here, the effects for non-refugees again flatten at around age ten, with the increasing pattern continuing for refugees. For both groups, the intermarriage probability conditional on being married goes down with age and flattens at around age 11 (panel d): those who arrive at age 15 have a 20-percentage-point lower probability of marrying a native than those that arrive at ages 0-3.

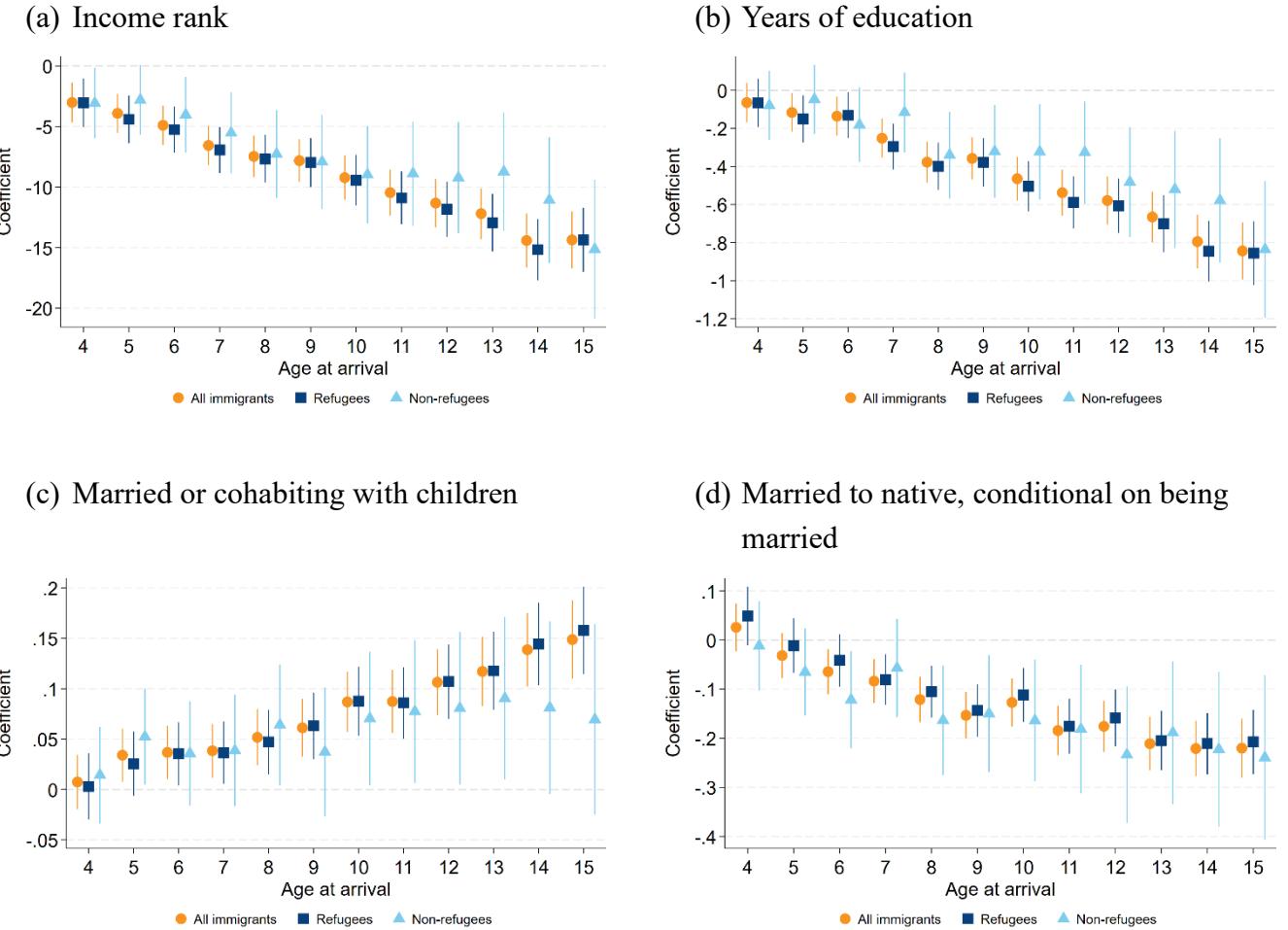
Overall, both groups experience lower income ranks, fewer years of education and reduced intermarriage rates when arriving later compared to arriving at ages 0-3. However, the effects flatten off only for non-refugees.

<sup>12</sup> In the refugee sample, 82.01% of sibling pairs have an age gap of at most 5 years; in the non-refugee sample, this percentage climbs slightly to 85.56%.

<sup>13</sup> We note that the sudden drops at age 15 are most likely driven by a small number of observations in that age category. Figure A.4 shows the analogous figures for when we pool ages at arrival 13-15 together in one category. While the drop is still there for the highest age category (Figure A.4a), it is much less abrupt.

Given that age at arrival matters for labor market, education, and intermarriage outcomes, our final step of inquiry is to estimate how much of the baseline effects of age at arrival on residential segregation can be explained by these three intermediate channels. We turn to this in the next section.

Figure 2: Effect of age at arrival on other integration outcomes



*Note:* The figure shows the  $\beta_a$  coefficients obtained when estimating equation (2) and their corresponding 95% confidence intervals.

*Source:* Own calculations on data from the GeoSweden database.

### 3.3 Decomposing the main effect on residential segregation

We decompose the effects of age at immigration on neighborhood integration into components attributable to labor market integration (through income rank and education) and social integration (through intermarriage) in the style of Heckman et al. (2013). While this exercise brings important insights into why we may observe the residential segregation patterns above, a word of caution is warranted with respect to this analysis. To be able to interpret these results as causal effects of the mediators, we need to make strong assumptions. In particular, we need to assume that all unobserved factors should be uncorrelated with both age at arrival and the mediators, and orthogonal to the link between the mediators and residential segregation.

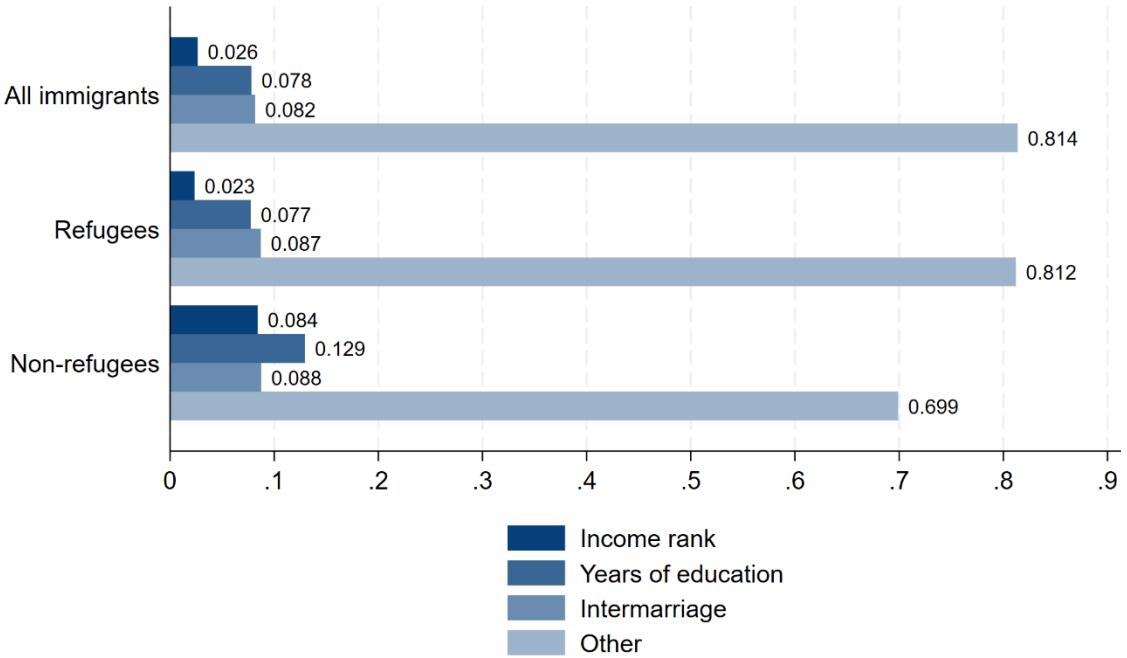
Additionally, we measure mediators and outcomes at the same age, raising potential reverse causality concerns. While this concern is minimal for education, which is largely completed before residential decisions, causality could run in both directions for income (e.g., neighborhoods affecting income through local job opportunities) and intermarriage (e.g., immigrants meeting spouses in their neighborhoods). For these reasons, we think of this method as describing patterns to help us better understand our results - showing which factors are associated with residential sorting patterns - rather than identifying strict causal mechanisms.

Since we are interested in how both labor market integration and social integration (through intermarriage) contribute to residential segregation, we conduct the decomposition analysis on the married sample. We describe in detail the steps involved in this exercise in section A.3. We estimate equation (2) with age of the child entering linearly in the decomposition exercise (that is, we decompose a linear effect of age at arrival). The main reason for this choice is clarity; instead of presenting a decomposition analysis for each and every age coefficient estimated in Figures 1-2, we present an overall decomposition analysis. We note, however, that our analysis has revealed non-linearities in effects and therefore these results should be interpreted with that caveat in mind.

Figure 3 shows the contributions of each channel to the overall effect on residential segregation. We see that the groups of refugees and non-refugees differ in how much each channel contributes. While for non-refugees income rank and years of education contribute roughly 20%, these two channels are only half as important for refugees. The intermarriage channel contributes equally in terms of absolute shares. Direct intermarriage effects were also found recently for Sweden (Jarvis et al. 2023), where rising intermarriage and cohabitation rates have also been observed over the recent decades (Elwert, 2020). Yet, given that for refugees there is a larger part of the variation that is unexplained, intermarriage actually contributes equally with respect to the other channels, whereas for non-refugees intermarriage is half as important. Figure A.5, which shows the results for the full sample, where we cannot estimate the contribution of intermarriage, also reveals that the unexplained part is larger for refugees than it is for non-refugees, suggesting the presence of other factors that prevent refugees from integrating residentially.

These results tie our previous findings together: while refugees and non-refugees arriving later integrate similarly in labor markets and marriage markets, they differ substantially in residential outcomes. For non-refugees, economic integration (income and education) explains a meaningful share of residential segregation, consistent with spatial assimilation theories where economic success facilitates residential mobility. For refugees, economic and social integration contribute more equally, but large unexplained residuals remain—particularly notable given their similar labor market and intermarriage patterns to non-refugees. This suggests that refugees face additional structural barriers to residential integration. Such barriers likely include housing market discrimination (Ahmed and Hammarstedt 2008; Molla et al. 2022) or dispersal policies that constrain initial settlement locations.

Figure 3: Decomposition



*Note:* The figure shows the contributions of income rank, years of education, intermarriage and a residual category to the overall effect on residential segregation in the siblings married sample. *Source:* Own calculations on data from the GeoSweden database.

#### 4 Conclusions

In this paper, we have shown that the age at which immigrant children—particularly those with refugee status—arrive in their new country significantly affects the level of segregation in their neighborhoods in adulthood. Our results indicate that early arrival can have a non-negligible contribution to the overall (municipality-level) segregation level. Our analysis of potential mechanisms tentatively suggests that economic factors play a larger role for non-refugees, whereas for refugees, intermarriage and economic variables contribute equally to explaining the variation in the effect of age at immigration.

These results suggest that integration policies should be differentiated both by age at arrival and by refugee status. Our results on the importance of economic integration channels indicate that policies strengthening school acclimatization, language acquisition, flexible schooling options, and labor market programs can help those arriving later in childhood. Interventions that foster social ties—such as programs facilitating contact with native peers—may be especially crucial for refugees, for whom intermarriage plays a larger role. More generally, policies need to recognize that late-arriving children require targeted educational and social support, and that refugees face unique constraints that extend beyond the labor market. Given

the importance of age at arrival, it is also worth noting the benefits of a fast asylum decision process, which allows early access to education and training for refugees.

While prior studies often lump together refugee- and non-refugee immigrants, we find different effect sizes, timing effects and channel importance with regard to segregation outcomes. How these results differ by ethnicity or country of origin, by different ages beyond 30 or for different cohorts would just be some ways to further this research on age at arrival and urban segregation. Finally, a closer inspection into the aggregation dynamics of different segregation measures (including entropy-based ones) could help to better uncover the micro-macro link behind general segregation dynamics.

## References

- Ahmed, Ali M., and Mats Hammarstedt. 2008. “Discrimination in the Rental Housing Market: A Field Experiment on the Internet.” *Journal of Urban Economics* 64 (2): 362–372.
- Aksoy, Cevat Giray, Panu Poutvaara, and Felicitas Schikora. 2023. “First Time Around: Local Conditions and Multi-Dimensional Integration of Refugees.” *Journal of Urban Economics* 137: 103588.
- Alexander, Rohan, and Zachary Ward. 2018. “Age at Arrival and Assimilation during the Age of Mass Migration.” *Journal of Economic History* 78(3): 904–37.
- Andersson, Roger, Åsa Bråmå, and Emma Holmqvist. 2010. “Counteracting Segregation: Swedish Policies and Experiences.” *Housing Studies* 25(2): 237–56.
- Andersson, Henrik, Sirus Dehdari, and Karl Oskar Lindgren. 2025. “Age at Migration and the Political Integration of Immigrants—Evidence from a Sibling Analysis.” *International Migration Review*. Advance online publication.
- Ansala, Laura, Ulla Hämäläinen, and Matti Sarvimäki. 2019. “Age at Arrival, Parents and Neighborhoods: Understanding the Educational Attainment of Immigrants’ Children.” *Journal of Economic Geography* 20(2): 459–80.
- Åslund, Olof, Anders Böhlmark, and Oskar Nordström Skans. 2015. “Childhood and Family Experiences and the Social Integration of Young Migrants.” *Labour Economics* 35: 135–44.
- Basu, Soma. 2018. “Age-of-Arrival Effects on the Education of Immigrant Children: A Sibling Study.” *Journal of Family and Economic Issues* 39(3): 474–93.
- Böhlmark, Anders. 2008. “Age at Immigration and School Performance: A Siblings Analysis Using Swedish Register Data.” *Labour Economics* 15(6): 1366–87.

- Brell, Courtney, Christian Dustmann, and Ian Preston. 2020. "The Labor Market Integration of Refugee Migrants in High-Income Countries." *Journal of Economic Perspectives* 34(1): 94–121.
- Cederström, Agneta, and Andrea Dunlavy. 2025. "A matter of measurement? A Swedish register-based study of migrant residential segregation and all-cause mortality." *SSM - Population Health* 30: 101793.
- Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor. 2008. "Is the Melting Pot Still Hot? Explaining the Resurgence of Immigrant Segregation." *Review of Economics and Statistics* 90(3): 478–97.
- Duncan, Otis Dudley, and Beverly Duncan. 1955. "Residential Distribution and Occupational Stratification." *American Journal of Sociology* 60(5): 493–503.
- Elwert, Annika. 2020. "Opposites Attract: Assortative Mating and Immigrant–Native Intermarriage in Contemporary Sweden." *European Journal of Population* 36(4): 675–709.
- Harder, Niklas, Lucila Figueroa, Rachel M. Gillum, Dominik Hangartner, David D. Laitin, and Jens Hainmueller. 2018. "Multidimensional Measure of Immigrant Integration." *Proceedings of the National Academy of Sciences* 115 (45): 11483–88.
- Heckman, James, Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review* 103(6): 2052–86.
- Helgesson, Magnus, et al. 2019. "Labour Market Marginalisation among Refugees from Different Countries of Birth: A Prospective Cohort Study on Refugees to Sweden." *Journal of Epidemiology & Community Health* 73(5): 407–15.
- Hermansen, Are Skeie. 2017. "Age at Arrival and Life Chances among Childhood Immigrants." *Demography* 54(1): 201–29.
- Jarvis, Benjamin F., Robert D. Mare, and Monica K. Nordvik. 2023. "Assortative Mating, Residential Choice, and Ethnic Segregation." *Research in Social Stratification and Mobility*. Advance online publication.
- Lemmermann, Dominique, and Regina T. Riphahn. 2018. "The Causal Effect of Age at Migration on Youth Educational Attainment." *Economics of Education Review* 63: 78–99.
- Malmberg, Bo, Eva K. Andersson, Michael M. Nielsen, and Karen Haandrikman. 2018. "Residential Segregation of European and Non-European Migrants in Sweden: 1990–2012." *European Journal of Population* 34(2): 169–93.
- Massey, Douglas S., and Nancy A. Denton. 1988. "The Dimensions of Residential Segregation." *Social Forces* 67(2): 281–315.

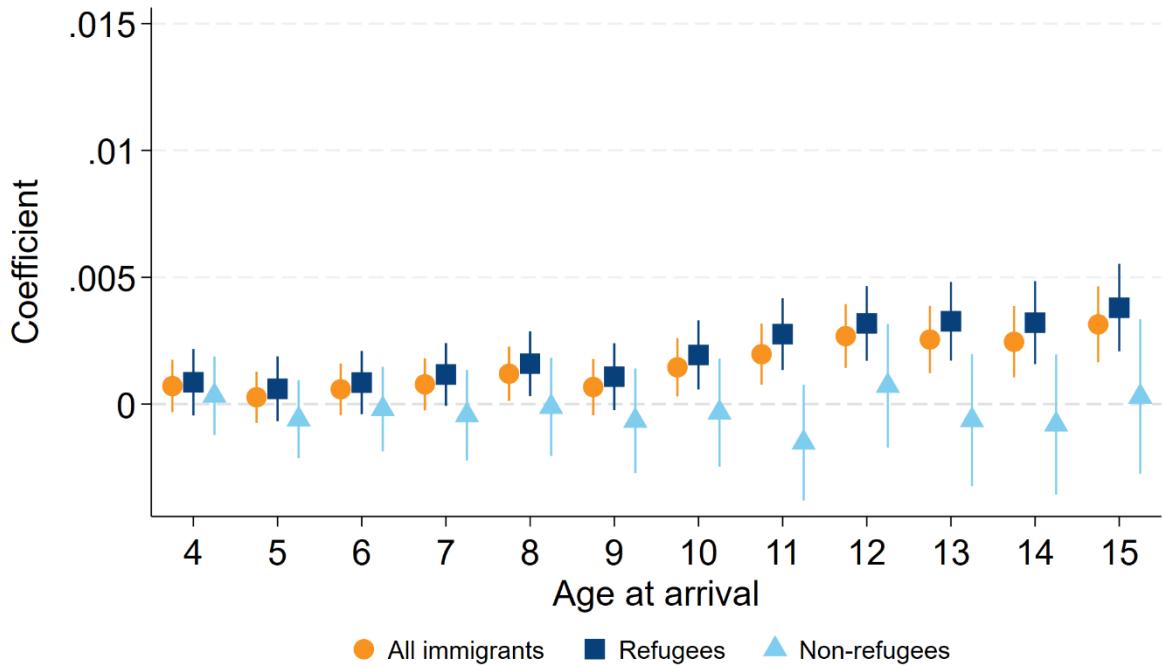
Molla, Hemrin, Caroline Rhawi, and Elina Lampi. 2022. “Name Matters! The Cost of Having a Foreign-Sounding Name in the Swedish Private Housing Market.” *PLOS ONE* 17 (6): e0268840.

Van den Berg, Gerard J., Petter Lundborg, Paul Nystedt, and Dan-Olof Rooth. 2014. “Critical Periods during Childhood and Adolescence.” *Journal of the European Economic Association* 12(6): 1521–57.

## A Appendix

### A.1

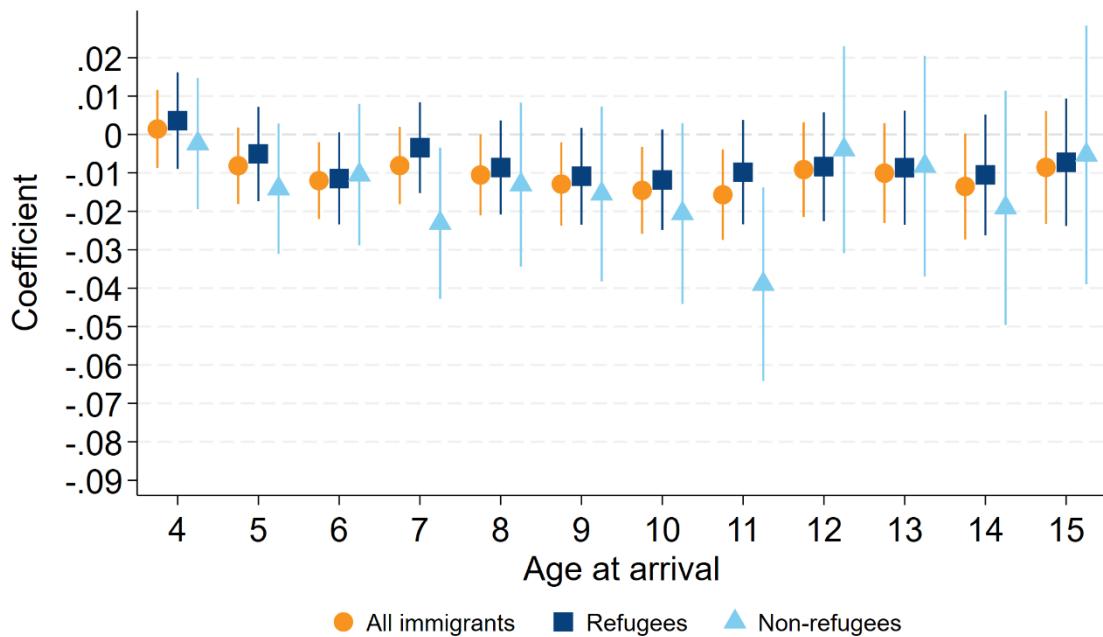
Figure A.1: Effect of age at arrival on the isolation index component



*Note:* The figure shows the  $\beta_a$  coefficients obtained when estimating equation (2) for the isolation index component and their corresponding 95% confidence intervals.

*Source:* Own calculations on data from the GeoSweden database.

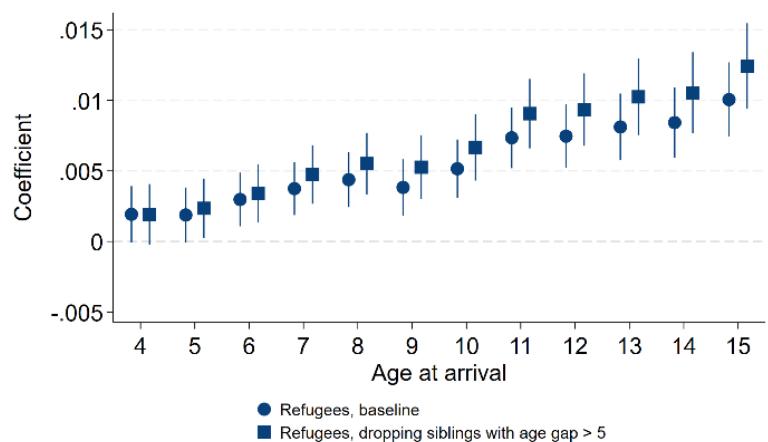
Figure A.2: Effect of age at arrival on the neighborhood share of immigrants



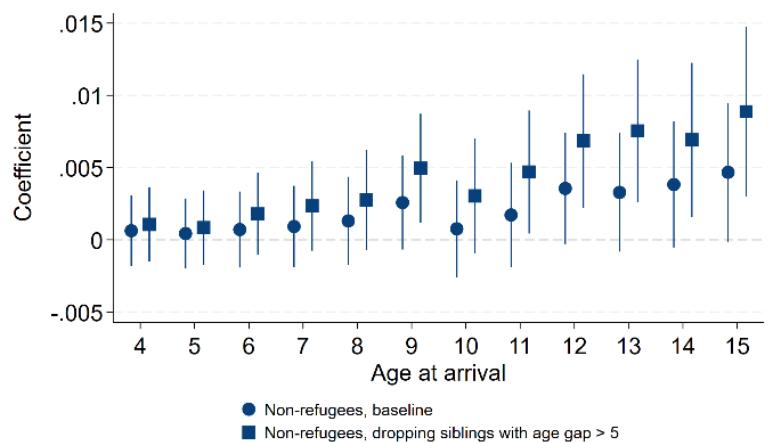
*Note:* The figure shows the  $\beta_a$  coefficients obtained when estimating equation (2) for the share of immigrants and their corresponding 95% confidence intervals.

*Source:* Own calculations on data from the GeoSweden database.

Figure A.3: Effect of age at arrival on the neighborhood contribution to dissimilarity index  
 (excluding siblings with large age gaps)

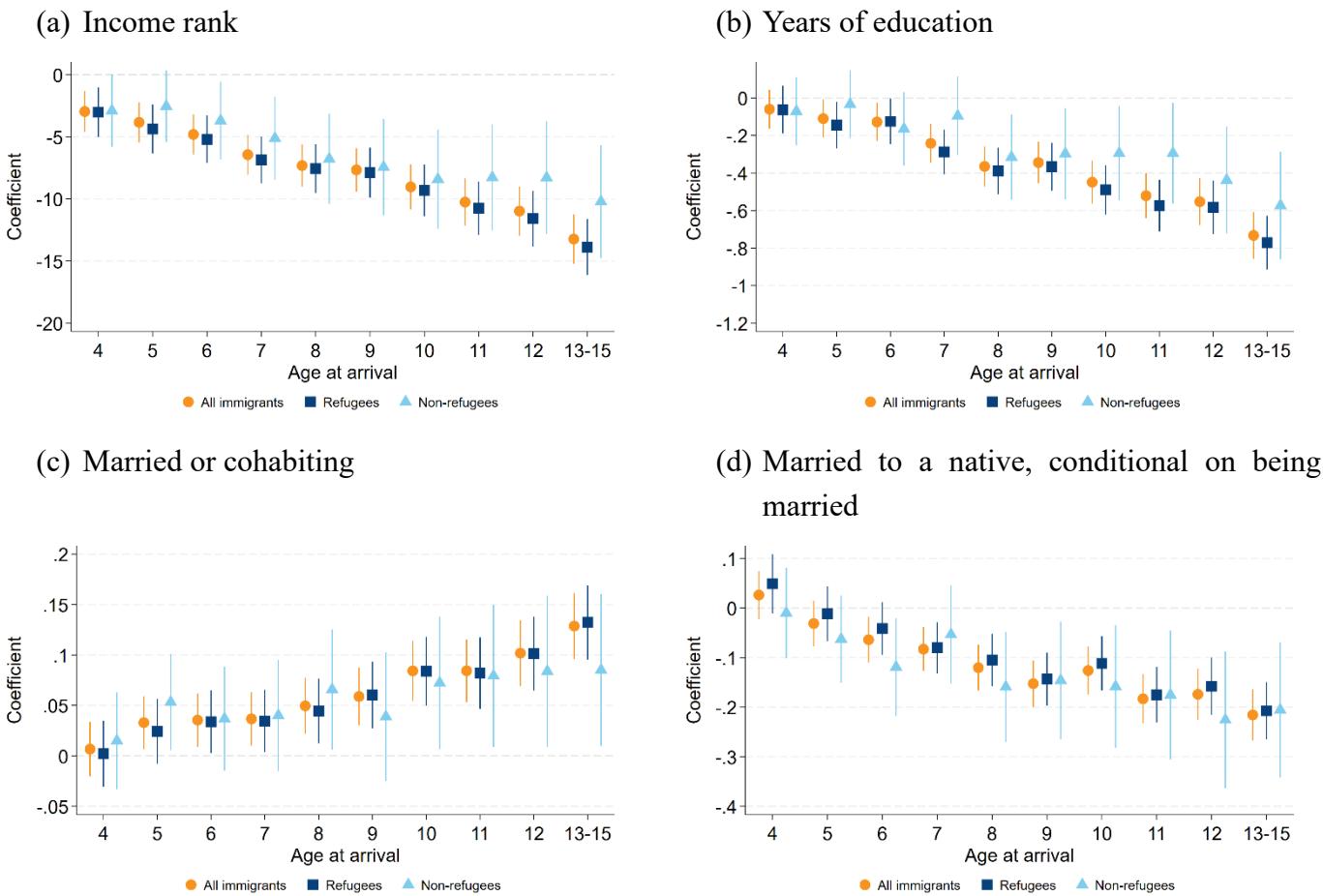


(a) Refugees



(b) Non-refugees

Figure A.4: Effect of age at arrival (alternative definition) on other integration outcomes



## A.2 Tables

Table A.1: Summary statistics for the full sample

	Mean	Std. dev.	No. of obs.
<i>Panel A: All immigrants</i>			
Neighborhood-level contribution to dissimilarity index (baseline)	0.019	0.031	82,135
Income rank	46.008	30.877	82,135
Years of education	12.396	2.285	81,364
Married	0.406	0.491	82,135
Intermarried	0.358	0.479	33,387
Female	0.469	0.499	82,135
First-born	0.579	0.494	82,135
Age at arrival	8.534	4.064	82,135
<i>Panel B: Refugees</i>			
Neighborhood-level contribution to dissimilarity index	0.019	0.032	56,494
Income rank	46.785	30.883	56,494
Years of education	12.448	2.293	56,024
Married	0.420	0.494	56,494
Intermarried	0.279	0.449	23,733
Female	0.465	0.499	56,494
First-born	0.512	0.500	56,494
Age at arrival	9.073	3.744	56,494
<i>Panel C: Non-refugees</i>			
Neighborhood-level contribution to dissimilarity index	0.018	0.029	25,641
Income rank	44.296	30.796	25,641
Years of education	12.282	2.264	25,340
Married	0.377	0.485	25,641
Intermarried	0.552	0.497	9,654
Female	0.479	0.500	25,641
First-born	0.726	0.446	25,641
Age at arrival	7.345	4.469	25,641

*Notes:* This table reports summary statistics for all immigrants, refugees and non-refugees in the full sample, respectively. Children are born between 1974 and 1987. We classify a child as a refugee if either their own permit is a refugee permit or, absent that information, if they have at least one parent classified as a refugee. The dissimilarity index is the absolute value of the individual component for each i-th neighborhood in equation 1. Intermarriage is marriage to a Swedish-born partner.

*Source:* Own calculations on data from the GeoSweden database.

Table A.2: Countries of origin, by refugee status

Country of origin	N, non- refugees	Share, non- refugees	N, refugees	Share, non- refugees
Denmark	356	0.034	0	0.000
Finland	867	0.082	0	0.000
Norway	456	0.043	0	0.000
Iceland	166	0.016	0	0.000
Former Yugoslavia	281	0.027	5512	0.143
Croatia	18	0.002	197	0.005
Slovenia	5	0.000	5	0.000
Bosnia	40	0.004	4196	0.109
Macedonia	15	0.001	117	0.003
Poland	804	0.076	489	0.013
Belgium	15	0.001	4	0.000
Romania	153	0.014	665	0.017
Czech Republic	72	0.007	94	0.002
Hungary	128	0.012	226	0.006
Greece	31	0.003	9	0.000
Great Britain	155	0.015	37	0.001
Ireland	4	0.000	0	0.000
Germany	261	0.025	114	0.003
France	15	0.001	23	0.001
Italy	20	0.002	18	0.000
Portugal	53	0.005	11	0.000
Netherlands	38	0.004	3	0.000
Austria	12	0.001	13	0.000
Switzerland	8	0.001	3	0.000
Bulgaria	28	0.003	301	0.008
Other small countries in Europe	3	0.000	77	0.002
Estonia	43	0.004	41	0.001
Latvia/Lithuania	24	0.002	3	0.000
Former Soviet Union	93	0.009	236	0.006
Russia	81	0.008	40	0.001
Ethiopia	77	0.007	768	0.020
Somalia	99	0.009	1389	0.036
Gambia	181	0.017	12	0.000
Tunisia	12	0.001	32	0.001
Morocco	66	0.006	16	0.000
Uganda	104	0.010	202	0.005
Algeria	16	0.002	33	0.001

Egypt	5	0.000	20	0.001
Eritrea	50	0.005	473	0.012
Other countries in Africa	326	0.031	401	0.010
Lebanon	632	0.060	3185	0.083
Syria	184	0.017	1912	0.050
Turkey	1,130	0.107	1698	0.044
Iraq	301	0.029	4054	0.106
Iran	334	0.032	6104	0.159
Other countries in West Asia	79	0.007	483	0.013
Vietnam	377	0.036	715	0.019
Thailand	513	0.049	9	0.000
China and Taiwan	77	0.007	29	0.001
Philippines	232	0.022	10	0.000
Japan	4	0.000	0	0.000
Afghanistan	32	0.003	458	0.012
Bangladesh	31	0.003	157	0.004
India	68	0.006	42	0.001
South Korea	11	0.001	0	0.000
Pakistan	89	0.008	62	0.002
Sri Lanka	41	0.004	106	0.003
Other countries in Asia	70	0.007	124	0.003
United States of America	49	0.005	25	0.001
Canada	16	0.002	7	0.000
Central America	149	0.014	408	0.011
Chile	411	0.039	2500	0.065
Bolivia	113	0.011	64	0.002
Peru	59	0.006	257	0.007
Brazil	41	0.004	11	0.000
Argentina	138	0.013	69	0.002
Colombia	28	0.003	119	0.003
Other countries in South America	115	0.011	30	0.001
Australia	10	0.001	0	0.000
Other countries in Oceania	15	0.001	0	0.000

*Note:* Country of origin by refugee status. Countries of origin with 3 or more observations shown.

*Source:* Own calculations on data from the GeoSweden database.

Table A.3: Baseline means

	All	Refugees	Non-refugees
<i>Panel A: Residential segregation outcomes</i>			
Neighborhood-level contribution to dissimilarity index	0.015	0.014	0.016
Dissimilarity index	0.693	0.695	0.691
Share immigrants in the neighborhood	0.340	0.357	0.319
<i>Panel B: Other integration outcomes</i>			
Income rank	47.895	49.619	45.739
Years of education	12.546	12.763	12.274
Married	0.350	0.330	0.375
Intermarried	0.534	0.475	0.600

*Note:* The baseline means refer to the pooled category of those who arrive between the ages of 0 and 3.

### A.3 Decomposition

The decomposition is conducted in three steps:

1. We first estimate equation (2) with a linear age variable and with the variables income rank, years of education and intermarriage as additional covariates, and save the coefficients on these three additional variables and the main effect of age. These coefficients are in columns (1)-(4) in Table A.4 for the married sample and columns (1)-(3) in Table A.5 for the full sample.
2. We then estimate equation (2) with a linear age at arrival variable, separately for each of the variables income rank, years of education and intermarriage (in the married sample only) as outcome variables. We save the coefficient on the age variable from each of these regressions (columns (5)-(7) in Table A.4 for the married sample and columns (4)-(5) in Table A.5 for the full sample).
3. Finally, we calculate the contribution of each of the three “channel” variables. This is done by multiplying the coefficient on each variable as estimated in the first step with the respective coefficient on age as estimated in the second step. This means that we weight the contribution of each variable to the main outcome by the effect of age on that variable. These estimated contributions can be found in columns (8)-(10) of Table A.4 for the married sample and columns (6)-(8) of Table A.5 for the full sample.

The total effect is equal to the main effect of age plus the contributions considered, and the shares are equal to each contribution divided by the total effect. These shares are presented in Panel A of Table A.3 and Figure 3 for the married sample; and in Panel B of Table A.3 and Figure A.5 for the full sample.<sup>14</sup>

---

<sup>14</sup> The decomposition presented in panel A of Table A.3 is based on those individuals that had married at age 30. The reason for this is that we want to decompose the main effects into all three intermediate channels. However, it can be noted that when we use the full sample and decompose the baseline effects into the labor market and education channels, we get shares for these intermediate channels that are very similar to those in panel A (see panel B).

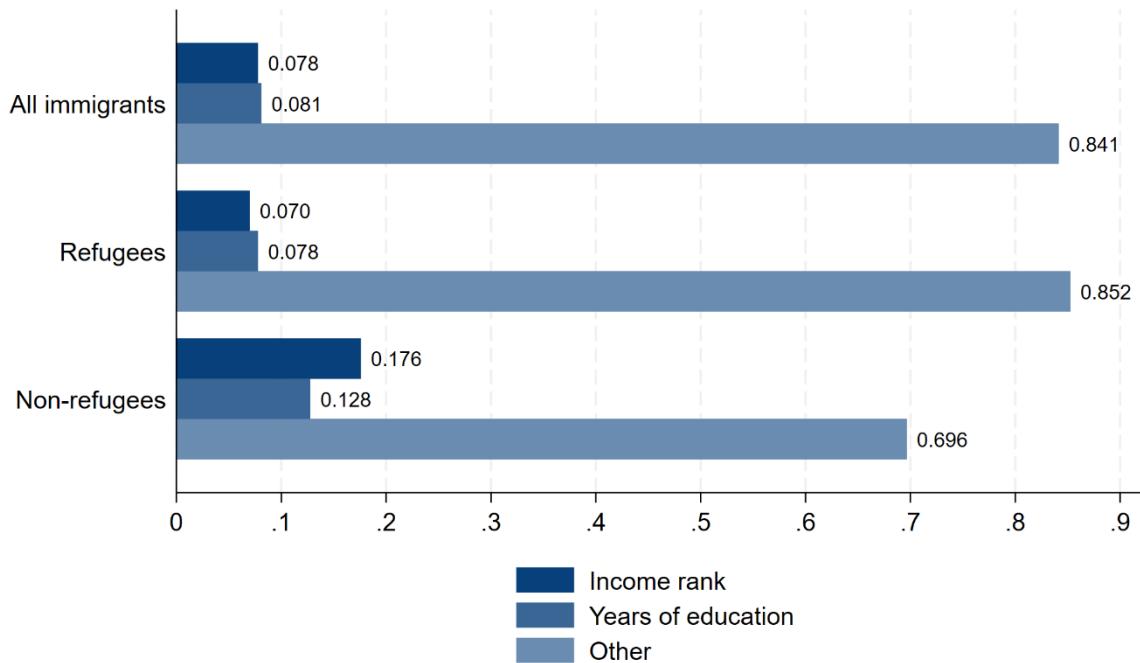
### A.3.1 Decomposition main results

Table A.4. Decomposition

	(1) All	(2) Refugees	(3) Non-refugees
<i>Panel A: Married sample</i>			
Income rank	0.026	0.023	0.084
Years of education	0.078	0.077	0.129
Intermarriage	0.082	0.087	0.088
Residual	0.814	0.812	0.699
<i>Panel B: Full sample</i>			
Income rank	0.078	0.070	0.176
Years of education	0.081	0.078	0.128
Residual	0.841	0.852	0.696

Note: Shares used to produce Figure 3 and Figure A.5.

Figure A.5: Decomposition, full siblings sample



Notes: The figure shows the contributions of income rank, years of education and a residual category to the overall effect on residential segregation in the siblings full sample.

Source: Own calculations on data from the GeoSweden data

### A.3.2 Steps to obtain decomposition shares

Table A.5 Decomposition; steps to obtain shares, married sample

	Coefficients from augmented eq. (1)				Effect of age on channels			Contributions				Shares			
	(1) Age	(2) I	(3) ED	(4) IM	(5) I	(6) ED	(7) IM	(8) I (2) × (5)	(9) ED (3) × (6)	(10) IM (4) × (7)	(11) T (1) + (8) + (9) + (10)	(12) I (8)/(11)	(13) ED (9)/(11)	(14) IM (10)/(11)	(15) R (1)/(11)
All immigrants	0.0007	0.0000	-0.0009	-0.0036	-1.0014	-0.0816	-0.0203	0.0000	0.0001	0.0001	0.0009	0.0265	0.0781	0.0817	0.8138
Refugees	0.0008	0.0000	-0.0009	-0.0040	-0.9855	-0.0840	-0.0204	0.0000	0.0001	0.0001	0.0009	0.0234	0.0774	0.0870	0.8121
Non-refugees	0.0003	0.0000	-0.0008	-0.0020	-1.1591	-0.0700	-0.0193	0.0000	0.0001	0.0000	0.0004	0.0841	0.1294	0.0875	0.6989

Table A.6 Decomposition; steps to obtain shares, full sample

	Coefficients from augmented eq. (1)				Effect of age on channels			Contributions				Shares		
	(1) Age	(2) I	(3) ED	(4) I	(5) ED	(6) I (2) × (4)	(7) ED (3) × (5)	(8) T (1) + (6) + (7)	(9) I (6)/(8)	(10) ED (7)/(8)	(11) R (1)/(8)			
All immigrants	0.0006	0.0000	-0.0008	-1.1516	-0.0709	0.0001	0.0001	0.0007	0.0777	0.0809	0.8414			
Refugees	0.0007	0.0000	-0.0008	-1.1715	-0.0733	0.0001	0.0001	0.0008	0.0698	0.0777	0.8525			
Non-refugees	0.0002	-0.0001	-0.0008	-1.0647	-0.0552	0.0001	0.0000	0.0003	0.1758	0.1277	0.6965			