Age at Arrival and Immigrant Segregation: A Between-Siblings Analysis *

Cristina Bratu† Matz Dahlberg‡ Sebastian Kohl§
Madhinee Valeyatheepillay¶

February 27, 2025

Abstract

Immigrant residential segregation has traditionally been high, particularly for refugees, though it has been declining over time. Understanding the factors that drive these patterns is crucial. While the literature on urban segregation has largely ignored the importance of when immigrants arrive, the age at arrival literature has neglected segregation as an outcome. This study bridges these gaps. Using variation in age at arrival between siblings, we show causally that immigrants who arrive earlier are more likely to live in less segregated neighborhoods as adults, with the effect being particularly strong for refugees. Switching from post- to pre-school age may lower total segregation by 6.6%. A decomposition analysis 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.

Keywords: Refugees, Immigrants, Ethnic segregation, Age at arrival, Intermarriage

*We thank Henrik Andersson, Yvonne Giesing, Stefano Lombardi, Panu Poutvaara and Janne Tukiainen for helpful comments, as well as participants at the 2020 Finnish Economic Association meeting in Tampere and Urban Lab seminar participants at Uppsala University. This is a heavily revised version of an earlier working paper that circulated under the title "Age at Arrival and Residential Integration" (CESifo WP No. 9181).

†VATT Institute for Economic Research, Finland; Uppsala Center for Fiscal Studies (UCFS); Uppsala Center for Labor Studies (UCLS). cristina.bratu@vatt.fi. Corresponding author.

‡Uppsala University, Sweden; Centre for Economic Studies (CESifo); UCFS; UCLS; Urban Lab at Uppsala University. matz.dahlberg@ibf.uu.se.

§Free University Berlin, sebastian.kohl@fu-berlin.de

¶Urban Lab at Uppsala University. m.valeyatheepillay@gmail.com

1 Introduction

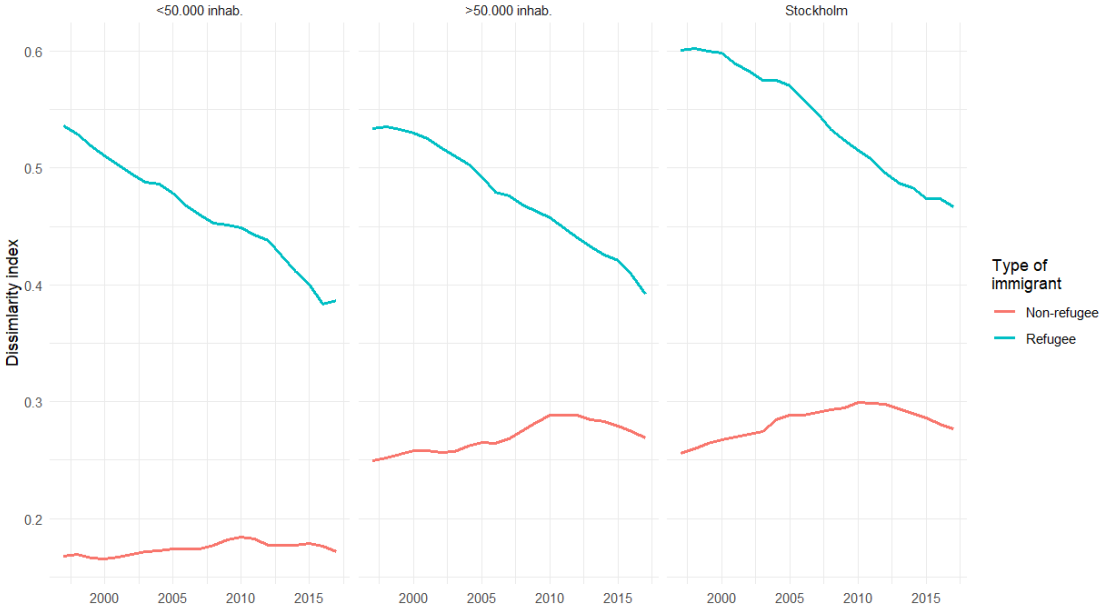
Ethnic segregation in European countries like Sweden increased rapidly in the post-WWII years, peaked in the late 20th century (Niedomysl et al., 2015; Friedrichs, 2008; Östh et al., 2014), and has since declined (Helbig and Jähnen, 2018).[1] Yet, even in Sweden, ethnic segregation remains a key topic in public debates, particularly in discussions on urban unrest (Malmberg et al., 2013) and police-designated vulnerable areas.

Among the explanations for this decline are the increasing absolute number of immigrants, the diminishing prevalence of ethnically concentrated neighborhoods that historically served as "anchor points" for new arrivals, and the expansion of anti-discrimination laws (Helbig and Jähnen, 2018; Cutler et al., 1997). However, one often overlooked factor in explaining these trends is the successful integration of immigrants themselves, particularly the role of age at arrival. Research shows that immigrants who arrive in early childhood tend to achieve better educational and labor market outcomes (Böhlmark, 2008; Hermansen, 2017; Alexander and Ward, 2018; Lemmermann and Riphahn, 2018; Ansala et al., 2019), enjoy better health (Van den Berg et al., 2014), and exhibit higher levels of social and political integration (Åslund et al., 2015; Andersson et al., 2025). While previous studies often use the minority share of a neighborhood as a measure of residential integration or attainment, the more complex outcome of residential segregation has largely been ignored.

In this paper, we therefore aim to bridge the immigrant integration and the urban segregation literatures by hypothesizing that earlier age at arrival may have contributed to the declining trends in ethnic segregation. We use high-quality register data from Sweden, where ethnic segregation – measured by the dissimilarity index within Sweden's 290 municipalities – has declined in recent decades, particularly among refugees (Figure 1). We apply a siblings design to cohorts born between 1974 and 1987 to study whether a younger age at arrival has an impact on residing in ethnically less segregated neighborhoods in adulthood (measured 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.

[1] This trend is similar to racial segregation in the United States (Massey and Tannen, 2015), though immigrant segregation in the U.S. reached an all-time high in 2000 (Cutler et al., 2008).

Figure 1: Ethnic segregation indices by immigrant groups and urbanization



Note: Population weighted-dissimilarity indices are computed for refugee and non-refugee immigrants separately. Trends are similar when using the Gini coefficient or the coefficient of variation.
Source: Own calculations on data from the GeoSweden database.

We use the neighborhood contribution to the dissimilarity index as a novel dependent variable and 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 unobservables may move abroad when their children are younger.[2] Our overall finding is that compared to immigrant children arriving between the ages of 0 and 3, immigrant children arriving later have a higher likelihood of residing in a segregated neighborhood. The likelihood is much higher for refugee than for non-refugee immigrants.

At first glance, this result may be at odds with the overall declining patterns we observe for the ethnic segregation of refugees. However, the aggregate picture is the result of two countervailing forces. On the one hand, arriving later results in worse residential outcomes for the cohorts born between 1974 and 1987. However, during the time period we study, refugees' age at arrival has gone down from an average of more than twelve to an average of around seven (Figure A.1). In addition, the younger cohorts are larger than the older cohorts. Hence, the overall segregation levels of refugees decreased dramatically between 2004 and 2017. A simple back-of-the-

[2]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.

envelope calculation estimates that switching immigration from post- to pre-school age could have lowered the initial 1997-segregation by about 6.6%.

We also probe into the different channels which would make migrants be more neighborhood-integrated by performing a decomposition analysis in the style of Heckman et al. (2013) to analyze how much of the effect of age at arrival on neighborhood integration goes through three important mechanisms identified in the literature: labor markets (earnings), education, and intermarriage, with the latter being defined as being married to or cohabiting (with children) with a Swedish-born partner. Our main decomposition finding is that intermarriage is a relatively more important channel for refugees than for non-refugees. In contrast, for non-refugees income and education are more important.

The main contribution of our paper is to integrate insights from the well-established age-at-arrival literature into the study of segregation while introducing residential segregation as a novel outcome within this framework. The study most closely related to ours is Åslund et al. (2015), but our work differs in two key ways. First, 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 non-Nordic European countries. Refugees remain an understudied group—largely due to data limitations—despite their distinct integration processes compared to other immigrant groups (see, e.g., Brell et al., 2020).

Second, 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.

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.[3] We measure the outcomes of interest at age 30 - an age by which residents become residentially settled - , hence we are implicitly restricting to those immigrants who do not return to their home country before that age. 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 that information, if they have at least one parent classified as a refugee. A non-refugee is an individual who does not fulfill these criteria. The "all immigrants" category pools together 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. To measure segregation, we primarily 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:

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

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 (Massey and Denton, 1988).

We use each individual's segregation index *component* as our outcome variable. This approach directly measures individual segregation contributions, as it takes into account the deviation of the destination neighborhood from city-wide averages. It also

[3]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.

allows us to infer the potential magnitude the individual contributions may have on total segregation.

Our neighborhood measure is the so-called DeSO, an administrative unit defined by Statistics Sweden such that the boundaries follow, to the extent possible, streets, waterways and railways. There are 5,984 DeSOs in Sweden, with population ranging from 700 to 2,700.

In equation 1, we define the group of immigrants either as born abroad to foreign-born parents or as born either in Sweden or abroad to foreign-born parents. That is, in the second definition, we include the second-generation in our immigrant definition.

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. We also show results using an alternative definition where intermarriage is defined as marriage with partners born in Sweden to Swedish parents only.

2.3 Empirical strategy

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

$$y_{ij} = \alpha + \sum_{a=4}^{15} \beta_a I(a_{ij} = a) + \mu \text{first-born}_{ij} + \theta \text{female}_{ij} + \phi_j + \eta_{ij} \tag{2}$$

where y_{ij} is the outcome of child i in family j, a_{ij} is the child's age at arrival in Sweden, ϕ_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 η_{ij} is the error term.[4] Those that arrive at ages 0-3 constitute the reference group.

Our empirical strategy addresses the concern that parents with better unobservables (in terms of, for example, motivation, parenting skills, and other variables that

[4]We consider siblings to have the same mother, when the mother is present in the registers; otherwise we use the information on the father.

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 β_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. 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: Summary statistics for the siblings sample

	Mean	Std. dev.	No. of obs.
Panel A: All immigrants			
Dissimilarity index (baseline)	0.02	0.03	48,962
Dissimilarity index (alternative)	0.02	0.02	48,962
Income rank	45.57	30.71	48,962
Years of education	12.30	2.27	48,547
Married	0.43	n/a	48,962
Intermarried (baseline)	0.31	n/a	20,817
Intermarried (alternative)	0.20	n/a	20,817
Female	0.47	n/a	48,962
First-born	0.38	n/a	48,962
Average age at arrival	8.72	3.78	48,962
Panel B: Refugees			
Dissimilarity index (baseline)	0.02	0.03	38,410
Dissimilarity index (alternative)	0.02	0.02	38,410
Income rank	46.08	30.73	38,410
Years of education	12.34	2.28	38,101
Married	0.43	n/a	38,410
Intermarried (baseline)	0.26	n/a	16,646
Intermarried (alternative)	0.16	n/a	16,646
Female	0.47	n/a	38,410
First-born	0.36	n/a	38,410
Average age at arrival	9.10	3.59	38,410
Panel C: Non-refugees			
Dissimilarity index (baseline)	0.02	0.03	10,552
Dissimilarity index (alternative)	0.01	0.02	10,552
Income rank	43.69	30.57	10,552
Years of education	12.14	2.24	10,446
Married	0.40	n/a	10,552
Intermarried (baseline)	0.51	n/a	4,171
Intermarried (alternative)	0.34	n/a	4,171
Female	0.48	n/a	10,552
First-born	0.43	n/a	10,552
Average age at arrival	7.37	4.13	10,552

Note: 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. In the baseline definition of the dissimilarity index, immigrants are all individuals born abroad to foreign-born parents. In the alternative definition, immigrants are all individuals born abroad to foreign-born parents *and* those born in Sweden to foreign-born parents. The baseline definition of intermarriage is marriage to a Swedish-born partner. The alternative version of intermarriage is marriage to a Swedish-born partner born to Swedish parents. Where standard deviations are not reported, the Mean column shows shares.
Source: Own calculations on data from the GeoSweden database.

Table 1 shows summary characteristics for each immigrant group in the siblings sample.[5] 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 when we use the baseline definition. We note a difference only when we use the alternative difference where we include the second-generation in the definition of immigrants. This suggests that refugees are more likely to be surrounded by people with immigrant backgrounds, regardless of country of birth. 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, especially when we restrict the definition of a native partner to consider only those born in Sweden to Swedish parents. On average, refugees arrive when they are three years older than non-refugees.

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, defined above as the neighborhood-level component (in absolute values) of the dissimilarity index of the neighborhood an individual lives in at age 30. 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. Finally, we decompose the main effect estimated in section 3.1 into parts attributable to the different channels in section 3.3.

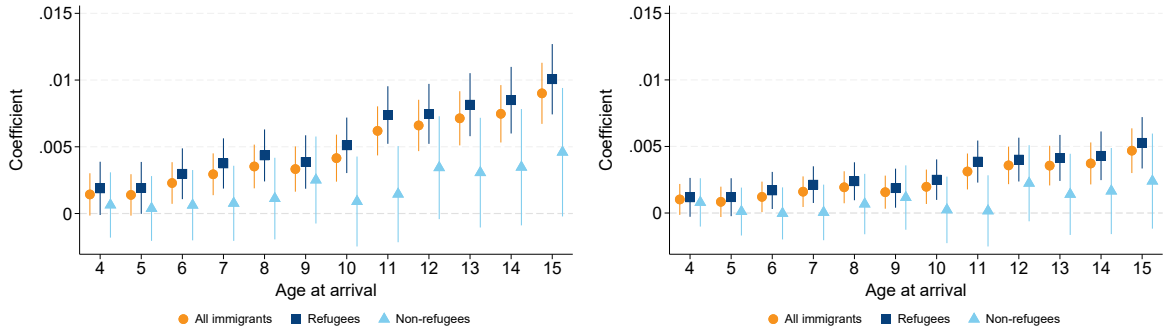
3.1 Effects on residential segregation

Figure 2 plots the β_a coefficients obtained when estimating equation (2) with the dissimilarity index component as the dependent variable. Focusing on Figure 2a, which uses the baseline dissimilarity index, we see that the later arrivals live in more segregated areas and that the effect increases roughly linearly with age. The effect on all immigrants (orange dots) is essentially entirely driven by the effect on refugees (dark blue squares). For non-refugees (light blue triangles), the effect is largely flat until the age of 11, when it starts increasing slightly. However, the effects are always smaller than for refugees, even among those that arrive at the age of 15. Figure 2b shows qualitatively similar patterns when we use the alternative dissimilarity index, but much smaller magnitudes. The results are very similar in direction and the different immigrant groups when using the individual contribution to the isolation index

[5] 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.

(see Appendix A.2) for the exposure instead of the evenness dimension of segregation.

Figure 2: Effect of age at arrival on dissimilarity index component



(a) Dissimilarity index (baseline)

(b) Dissimilarity index (alternative)

Note: The figure shows the β_a coefficients obtained when estimating equation (2) and their corresponding 95% confidence intervals.
Source: Own calculations on data from the GeoSweden database.

The absolute effect sizes appear obviously very small, as they refer to each individual's contribution to an index which ranges between 0 and 1. To get a more intuitive idea of how age at arrival effects on the dissimilarity neighborhood component contribute to total segregation in a municipality, we additionally estimate a model using a dummy variable for arriving before or after the age of seven instead of the categorical age dummies.[6] This allows for a back-of-the-envelope calculation, where we plug the coefficients of the age dummy into the dissimilarity index formula and assume the average neighborhood number of 20. Using the initial 1997 index values of 0.27, 0.54 and 0.19 for immigrants, refugees and non-refugees (cf. Figure 1), respectively, we arrive at respective segregation increases of 6.6%, 3.9% and 3.3% for a switch from pre-school age of arrival to school age (or later), which is not a negligible magnitude. This needs to be read as a broad average value which ignores the considerable differences between rural and the larger urban municipalities, which include more neighborhoods and higher segregation levels to start with, particularly for refugee immigrants.

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 integra-

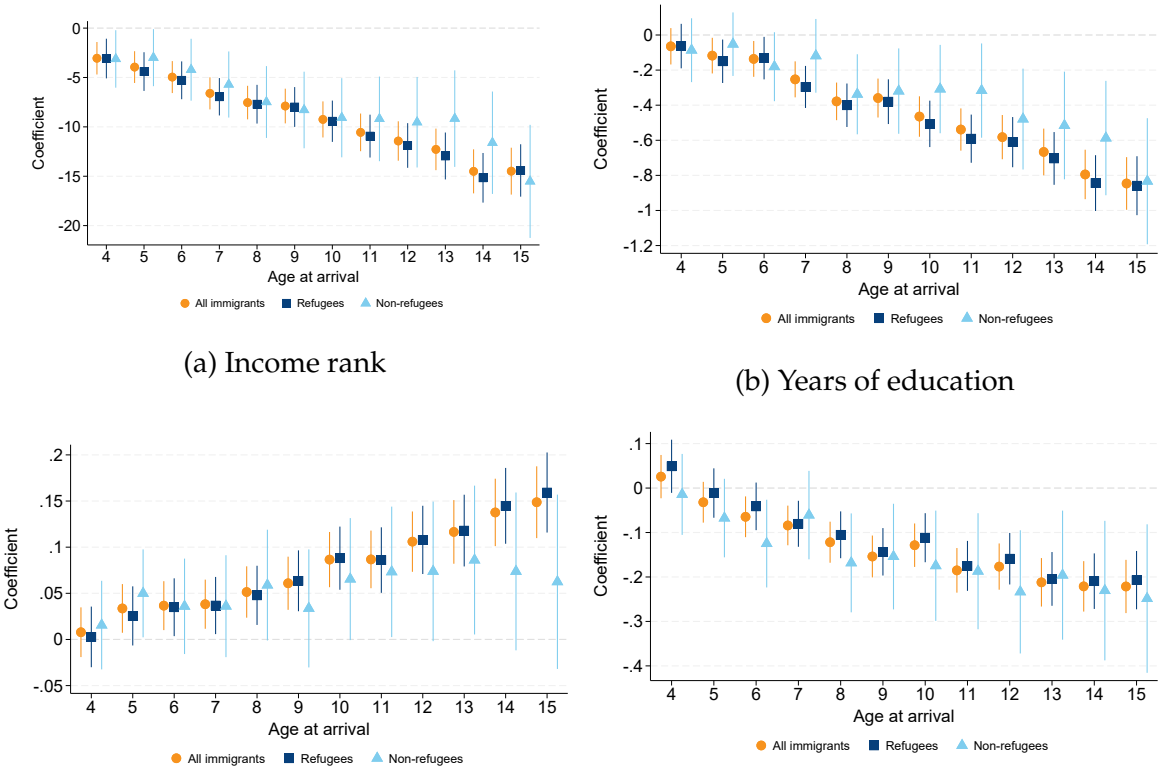
[6]This analysis yields a coefficient of 0.0018131 for all immigrants, 0.0021131 for refugees and 0.0006381 for non-refugees.

tion. 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 that are roughly similar in magnitude for both refugees and non-refugees. For example, arriving at age fifteen compared to ages 0-3 results in being ranked fifteen percentile ranks lower in the national income distribution as well as .8 fewer years of education. On the other hand, the probability of being married at age thirty increases with age at arrival. Here, the patterns for refugees and non-refugees differ slightly, in that the effects flatten for non-refugees at around age ten, but the increasing pattern continues for refugees. For both groups, the intermarriage probability conditional on being married goes down with age: those that arrive at age fifteen have a 20 percentage point lower probability of marrying a native than those that arrive at ages 0-3.

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 integration can be explained by these three intermediate channels. We turn to this in the next section.

Figure 3: Effect of age at arrival on other integration outcomes



(a) Income rank

(b) Years of education

(c) Married or cohabiting

(d) Married to native, conditional on being married

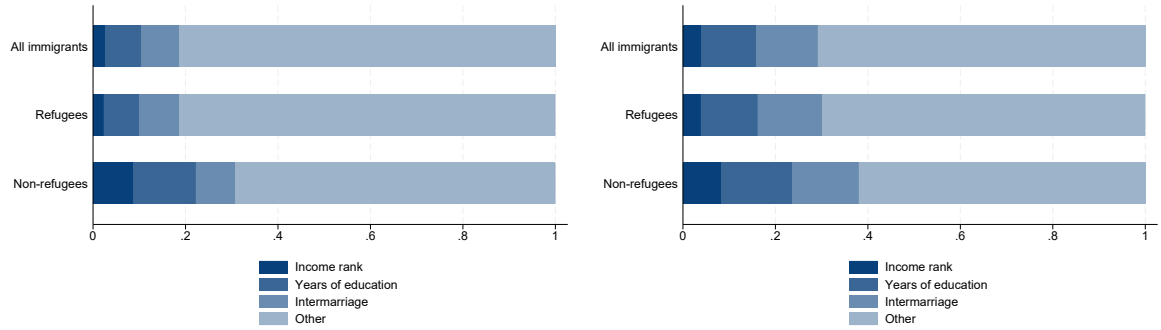
Note: The figure shows the β_a coefficients obtained when estimating equation (2) and their corresponding 95% confidence intervals.
Source: Own calculations on data from the GeoSweden database.

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. We describe in detail the steps involved in this exercise in section A.3. By necessity, we have to restrict to the sample of married individuals. We show the results for the full sample, where we cannot estimate the contribution of intermarriage, in figure A.3.

Since the estimated effects observed in Figures 2-3 are fairly linear, 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).[7]

Figure 4: Decomposition



(a) Dissimilarity index (baseline) (b) Dissimilarity index (alternative)

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

Figure 4a shows the contributions of each channel to the overall effect on residential segregation, for the baseline definition of 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, but 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 4b shows similar patterns, but larger magnitudes.

While this exercise brings important insights into why we may be observing the residential segregation patterns above, a word of caution is warranted with respect to

[7] Another reason for this choice is clarity; instead of presenting a decomposition analysis for each and every age coefficient estimated in Figures 2-3, we present an overall decomposition analysis.

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. For this reason, we think of this method as rather a descriptive tool to help us better understand our results.

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 ethnic segregation in their neighborhoods in adulthood. Our results indicate that early arrival can have a non-negligible contribution to the overall ethnic segregation level. Our analysis of potential mechanisms 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.

Our findings regarding the importance of different channels resonates well with micro-studies investigating the direct effect of inter-ethnic marriage choice and their residential choice on segregation in the US (Iceland and Nelson, 2010; Gabriel, 2015; Gabriel and Spring, 2019), where interracial marriages have been rising. Direct 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). Intermarriage has previously been identified as an important mechanism of social integration in other domains than residential integration (e.g. labor markets), most notably in the United States (Villazor, 2017).

Residential integration is generally measured by the share of different ethnic and racial minorities. Instead, this study uses the individual-level components of widely used segregation indices as an outcome variable, thus bridging the gap between the age at arrival literature and the urban segregation literature. Our interpretation of magnitudes suggests that the age-at-arrival effects are not negligible.

Finally, the effects we find are strongest for refugees when compared to non-refugee immigrants. These two groups, often lumped together in prior studies, do not only differ descriptively in terms of segregation levels. Their age at arrival is also much less important for predicting how segregated the neighborhoods of adult non-refugees eventually are and how important different channels of integration are, intermarriage in particular. How these results differ by ethnicity or country of origin would just be one way to further this research on age at arrival and urban segregation.

References

Alexander, Rohan and Zachary Ward (2018). "Age at arrival and assimilation during the age of mass migration". *The Journal of Economic History* 78.3, 904–937.

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*, 01979183241309579.

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, 459–480.

Åslund, Olof, Anders Böhlmark, and Oskar Nordström Skans (2015). "Childhood and family experiences and the social integration of young migrants". *Labour Economics* 35, 135–144.

Böhlmark, Anders (2008). "Age at immigration and school performance: A siblings analysis using Swedish register data". *Labour Economics* 15.6, 1366–1387.

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.

Cutler, Edward Glaeser, and Vigdor (1997). "The Rise and Decline of the American Ghetto". *NBER Working Paper* No. 5881.

Cutler, David M, Edward L Glaeser, and Jacob L Vigdor (2008). "Is the melting pot still hot? Explaining the resurgence of immigrant segregation". *The Review of Economics and Statistics* 90.3, 478–497.

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.

Friedrichs, Jürgen (2008). *Gespaltene Städte? Soziale und ethnische Segregation in deutschen Großstädten*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Gabriel, Ryan (2015). "A Middle Ground? Residential Mobility and Attainment of Mixed-Race Couples". *Demography* 53.1, 165–188.

Gabriel, Ryan and Amy Spring (2019). "Neighborhood Diversity, Neighborhood Affluence: An Analysis of the Neighborhood Destination Choices of Mixed-Race Couples With Children". *Demography* 56.3, 1051–1073.

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.

Helbig, Marcel and Stefanie Jähnen (2018). "Wie brüchig ist die soziale Architektur unserer Städte? Trends und Analysen der Segregation in 74 deutschen Städten". Report. WZB Discussion Paper.

Hermansen, Are Skeie (2017). "Age at arrival and life chances among childhood immigrants". *Demography* 54.1, 201–229.

Iceland, John and Kyle Anne Nelson (2010). "The residential segregation of mixed-nativity married couples". *Demography* 47.4, 869–893.

Jarvis, Benjamin F, Robert D Mare, and Monica K Nordvik (2023). "Assortative mating, residential choice, and ethnic segregation". *Research in Social Stratification and Mobility*, Online first.

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 Andersson, and John Östh (2013). "Segregation and urban unrest in Sweden". *Urban geography* 34.7, 1031–1046.

Massey, Douglas S. and Nancy A. Denton (1988). "The Dimensions of Residential Segregation". *Social Forces* 67.2, 281–315.

Massey, Douglas S and Jonathan Tannen (2015). "A research note on trends in black hypersegregation". *Demography* 52.3, 1025–1034.

Niedomysl, Thomas, John Östh, and Jan Amcoff (2015). *Boendesegregationen i Skåne [Residential segregation in Skåne]*. Kristianstad: Region Skåne.

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, 1521–1557.

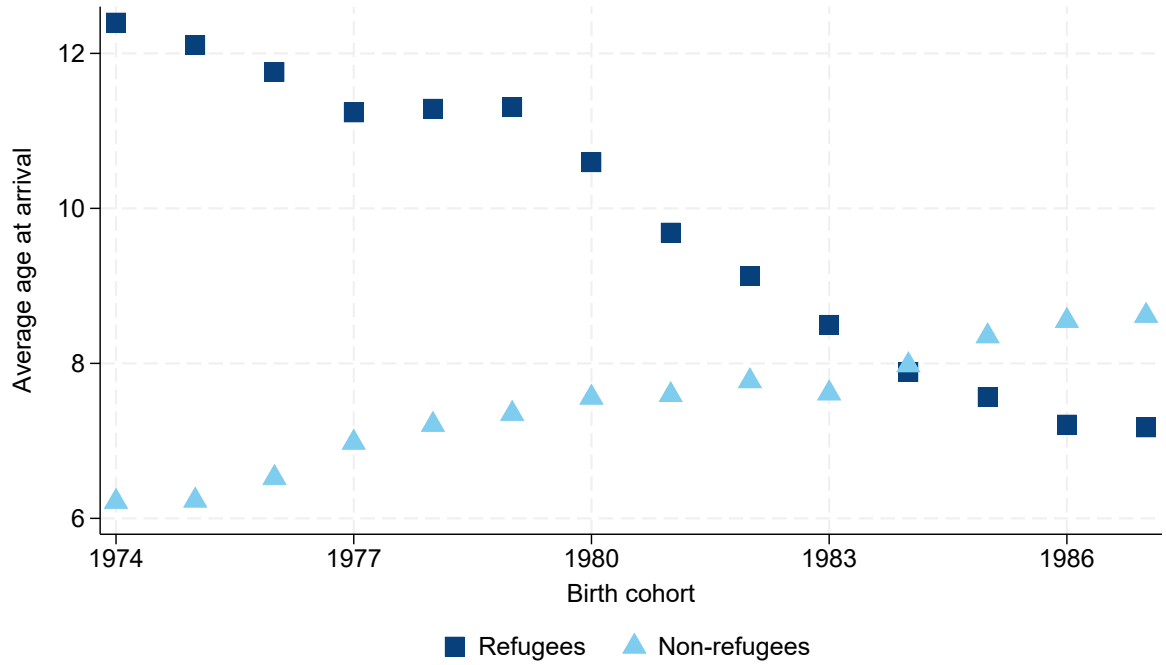
Villazor, Rose Cuison (2017). "Residential segregation and interracial marriages". *Fordham L. Rev.* 86, 2717.

Östh, John, Jan Amcoff, and Thomas Niedomysl (2014). "Segregation i Stockholmsregionen. Kartläggning med EquiPop [Segregation in the Stockholm region. Mapping with EquiPop]". Demografisk rapport 2014: 09. Stockholm: Stockholms Läns Landsting.

A Appendix

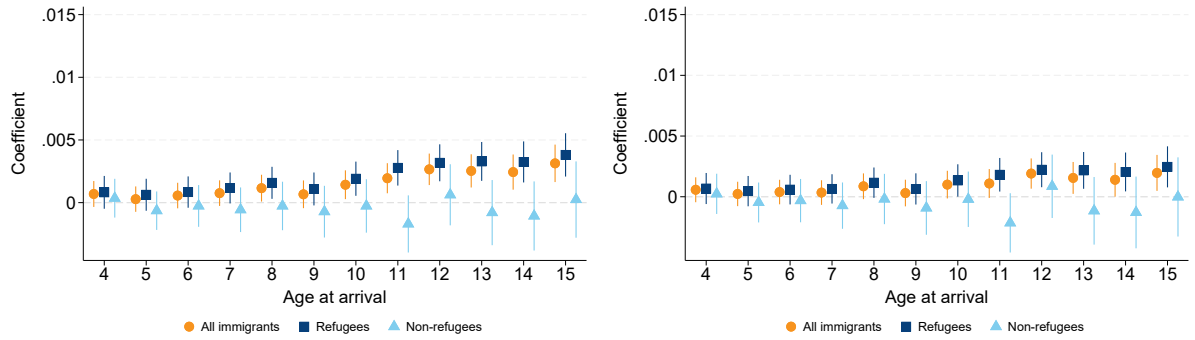
A.1 Figures

Figure A.1: Average age at arrival by birth cohort



Note: The figure shows average age at arrival by birth cohort, separately for refugees and non-refugees.
Source: Own calculations on data from the GeoSweden database.

Figure A.2: Effect of age at arrival on the isolation index component



Note: The figure shows the β_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.

A.2 Tables

Table A.1: Summary statistics for the full sample

	Mean	Std. dev.	No. of obs.
Panel A: All immigrants			
Dissimilarity index (baseline)	0.02	0.03	82,113
Dissimilarity index (alternative)	0.01	0.02	82,113
Income rank	46.01	30.87	82,113
Years of education	12.40	2.28	81,342
Married	0.41	n/a	82,113
Intermarried (baseline)	0.36	n/a	33,387
Intermarried (alternative)	0.24	n/a	33,387
Female	0.47	n/a	82,113
First-born	0.58	n/a	82,113
Average age at arrival	8.53	4.06	82,113
Panel B: Refugees			
Dissimilarity index (baseline)	0.02	0.03	56,477
Dissimilarity index (alternative)	0.02	0.02	56,477
Income rank	46.80	30.87	56,477
Years of education	12.45	2.29	56,009
Married	0.42	n/a	56,477
Intermarried (baseline)	0.28	n/a	23,732
Intermarried (alternative)	0.17	n/a	23,732
Female	0.46	n/a	56,477
First-born	0.51	n/a	56,477
Average age at arrival	9.07	3.74	56,477
Panel C: Non-refugees			
Dissimilarity index (baseline)	0.02	0.03	25,636
Dissimilarity index (alternative)	0.01	0.02	25,636
Income rank	44.29	30.80	25,636
Years of education	12.28	2.26	25,333
Married	0.38	n/a	25,636
Intermarried (baseline)	0.55	n/a	9,655
Intermarried (alternative)	0.39	n/a	9,655
Female	0.48	n/a	25,636
First-born	0.72	n/a	25,636
Average age at arrival	7.34	4.47	25,636

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. In the baseline definition of the dissimilarity index, immigrants are all individuals born abroad to foreign-born parents. In the alternative definition, immigrants are all individuals born abroad to foreign-born parents *and* those born in Sweden to foreign-born parents. The baseline definition of intermarriage is marriage to a Swedish-born partner. The alternative version of intermarriage is marriage to a Swedish-born partner born to Swedish parents. Where standard deviations are not reported, the Mean column shows shares.

Table A.2: Baseline means

	All immigrants	Refugees	Non-refugees
Panel A: Residential integration outcomes			
Dissimilarity index (baseline)	0.015	0.014	0.016
Dissimilarity index (alternative)	0.012	0.011	0.013
Panel B: Other integration outcomes			
Income rank	47.935	49.667	45.770
Years of education	12.549	12.768	12.275
Married	0.350	0.330	0.375
Intermarried (baseline)	0.533	0.475	0.598
Intermarried (alternative)	0.314	0.254	0.381

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.

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 as outcome variables. We save the coefficient on the age variable from each of these regressions (columns (5)-(7) in Table A.4).

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.

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 Table A.3 and Figure 4a for the married sample.[8]

[8]The decomposition presented in 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 Table A.3, see Table A.5 and the corresponding Table A.6 with the estimates obtained at steps 1-3 in the decomposition exercise.

A.3.1 Married sample

Table A.3: Decomposition

	Income rank	Years of education	Intermarriage	Residual
All immigrants	0.0266	0.0783	0.0814	0.8137
Refugees	0.0233	0.0771	0.0866	0.8129
Non-refugees	0.0875	0.1351	0.0853	0.6921

Table A.4: 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	-1.0083	-0.0819	-0.0204	0.0000	0.0001	0.0001	0.0009	0.0266	0.0783	0.0814	0.8137
Refugees	0.0008	0.0000	-0.0009	-0.0040	-0.9816	-0.0838	-0.0204	0.0000	0.0001	0.0001	0.0009	0.0233	0.0771	0.0866	0.8129
Non-refugees	0.0003	0.0000	-0.0008	-0.0019	-1.2204	-0.0725	-0.0199	0.0000	0.0001	0.0000	0.0004	0.0875	0.1351	0.0853	0.6921

A.3.2 Full sample

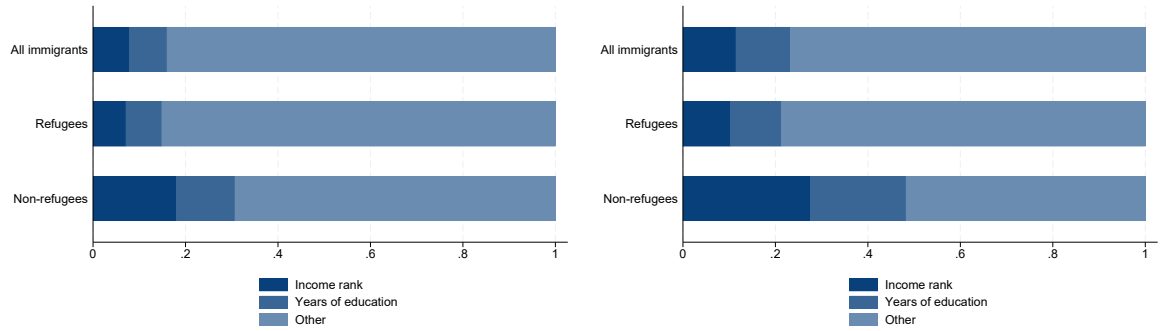
Table A.5: Decomposition

	Income rank	Years of education	Residual
All immigrants	0.0789	0.0809	0.8403
Refugees	0.0708	0.0776	0.8516
Non-refugees	0.1799	0.1268	0.6933

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 $(2) \times (4)$	(7) ED $(3) \times (5)$	(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	-1.1599	-0.0710	0.0001	0.0001	0.0007	0.0789	0.0809	0.8403
Refugees	0.0007	0.0000	-0.0008	-1.1732	-0.0736	0.0001	0.0001	0.0008	0.0708	0.0776	0.8516
Non-refugees	0.0002	-0.0001	-0.0008	-1.1016	-0.0545	0.0001	0.0000	0.0003	0.1799	0.1268	0.6933

Figure A.3: Decomposition



(a) Dissimilarity index (baseline) (b) Dissimilarity index (alternative)

Notes: The figure shows the contributions of income rank, years of education and a residual category to the overall effect on residential segregation.
Source: Own calculations on data from the GeoSweden database.