The impact of childhood immigrant exposure on adult intermarriage*

AINO KALMBACH
Aalto University School of Business and Helsinki GSE

MATTI SARVIMÄKI
Aalto University School of Business and Helsinki GSE

LIAM WREN-LEWIS
Paris School of Economics and INRAE

March 6, 2022

Abstract

This paper examines the impact of exposure to immigrants during childhood on natives' marriage behavior when they are adults. We use extremely high-resolution spatial data on where everyone in Finland born between 1977 and 1999 grew up to calculate the share of immigrants among each individual's immediate neighborhood, and then use naturally exogenous across-cohort within-location variation in immigrant shares to examine the impact of childhood exposure. We show that greater immigrant contact as a child significantly increases the probability that a native will marry an immigrant as an adult. Further results suggest that changes in attitudes or preferences are likely to drive at least part of this result.

*We thank Pasi Saukkonen for helpful comments. We gratefully acknowledge funding from CEPREMAP, the Institut Convergences Migration and the ANR (ANR-20-CE41-0014-01) which supported this work.

1. Introduction

Immigrant integration and discrimination towards migrants are increasingly salient issues across many Western countries. Intermarriage rates between natives and immigrants have long been considered a useful indicator of inter-ethnic relations, partly because they simultaneously capture the results of both immigrant behavior and the behavior of natives towards them (Kalmijn, 1998; Meng and Gregory, 2005). Yet, while research has shed light on what makes *immigrants* more likely to intermarry, little is known about what determines *natives*' choices to interact with and ultimately marry immigrants (Furtado and Trejo, 2013).

Social interaction has long been postulated as a potential means of reducing prejudices (e.g. Allport, 1954). Yet the potential for contact between immigrants and native population to change attitudes is unclear, with impacts varying across contexts and types of contact (e.g. Dustmann et al., 2019; Steinmayr, 2021). Moreover, while recent research is teaching us more about how contact can reduce prejudices, little evidence exists on whether the impacts can be long-term or substantial enough to impact decisions such as marriage (Paolini et al., 2021). An important question, therefore, remains relatively open: does contact with immigrants have long term impacts on natives' behavior?

In this paper, we aim to answer this question by investigating whether native's marriage decisions are influenced by contact they had with immigrants as a child. To do so, we use novel high-resolution Finnish registry data which allows us to construct for each individual a small cohort of children that they are likely to have interacted with based on their birth year, gender, and precise geographical location. We then exploit quasi-random variation in the immigrant share within locations across birth years to estimate the impact of childhood immigrant exposure.

Our main contribution is to show that natives who are exposed to more immigrants as children are more likely to marry an immigrant as an adult. This effect is driven by peers of the same birth cohort and same-sex, meaning that it is not driven by a larger pool of potential spouses, but instead is likely the result of friendships made during childhood. The effect is comparable for both men and women, and is of a sizable magnitude: Increasing the immigrant share of a child's neighborhood by one standard deviation increases their probability of marrying an immigrant by 10 percent of the mean.

We then go further by investigating impacts on other outcomes to gain insights into the mechanisms behind this result. First, we show that there are no similar impacts of an individual's exposure on educational performance, employment, or the share of immigrants among an individual's colleagues or neighbors as an adult. This implies that the result is not being driven by greater exposure leading individuals' to meet more potential immigrant partners through their residential location or workplace. Moreover, we also show that individuals exposed to more migrants as a child are significantly more likely to marry an immigrant who has never lived near to the location where the individual grew up. This suggests that the impact we find is not entirely driven by direct changes in individuals' social networks based on their childhood friends. Finally, we find an effect from immigrants living nearby even if they did not attend the same middle-school or high-school, suggesting that contact with individuals outside of these domains is also important for changing behavior.

This paper adds to an emerging literature on the long-run impacts of contact with ethnic minorities on behavior. In the US, Billings et al. (2021) and Brown et al. (2021) show that white children with greater exposure to blacks are more likely to register as Democrats as adults, while Bursztyn et al. (2021) shows that individuals living in counties with higher shares of Arab-Muslim migrants are less prejudiced towards this group and donate more to charities in Arab countries. Similarly, in the UK, Schindler and Westcott (2021) find less racial prejudice in areas where African American soldiers were posted seventy years previously. We complement these findings by exploiting a different source of variation in ethnic minority exposure and looking at an outcome which helps us understand the breadth of change in inter-ethnic relations.

Perhaps the closest article to ours is Merlino et al. (2019), who use a US panel to look at whether white adults are more likely to cohabit with blacks if they had a greater share of blacks in their secondary school cohort.[1] Although the context of inter-ethnic relations in the US is very different from that between immigrant groups and natives in Finland, they find results that are consistent with ours, i.e., greater childhood exposure to blacks of the same sex increases interracial cohabitation. Beyond the very different context, this paper has several advances, including taking advantage of the much more comprehensive data to show that the results are not driven by attrition, selection into schools, or location-specific time-trends.[2]

[1] A recent working paper by Holmlund et al. (2021) also looks at long-term impacts of exposure to immigrants in schools among four cohorts of children in Sweden. They find non-linear impacts on outcomes including the probability of having any children and having a child with an immigrant. They do not look at impacts on the probability of marrying or living with an immigrant partner.

[2] Also related is the paper of Bazzi et al. (2019), who look at the impact of interethnic mixing on intermarriage by exploiting a natural experiment in Indonesia. Consistent with our results, they find that there are significant impacts of ethnic composition on intermarriage, but translating results across contexts is made difficult since interethnic mixing is community-wide and persistent, and in most of the villages they study there is not a dominant ethnicity to the extent whites are in Finland.

2. Context

Immigration to Finland was heavily restricted until the 1990s and the foreign-born population remained minuscule. In 1990, only 1.3% percent of the population was foreign-born—and roughly half of them were offspring of Finnish emigrants, while the rest were largely Western Europeans with a Finnish spouse. Only 0.04% of the population was born in Africa and 0.1% in Asia in 1990.[3] Finland's immigration policy then slowly became more accomodating and a relatively large number of immigrants from neighboring Russia and Estonia moved to Finland during the 1990s. In addition, more refugees were admitted, first primarily from former Yugoslavia, Somalia and Turkey, and later also from Iraq and Afghanistan. Nevertheless, in comparison to most Western European countries, Finland's immigrant population remains relatively small with 7.6% of the population being foreign-born in 2020.

In our analysis, we follow Statistics Finland's definition of 'foreign background', i.e., a person is categorized as an immigrant if neither of her parents was born in Finland. Furthermore, we restrict our attention to the groups that face the most discrimination and prejudice. Here, we draw from the literature on 'ethnic hierarchies' based on surveys inquiring natives about their attitudes towards different immigrant groups. In these surveys, Black African, Turkish, Kurdish, Russian, Arab, and Somali immigrants consistently show up at the bottom of the hierarchy, while attitudes towards European, American, and Japanese immigrants are the most favorable (Jaakkola, 2000, 2005). The survey evidence is also corroborated by experimental evidence by Ahmad (2020), who finds that fictitious job applications from Somali and Iraqi background candidates have response rates of 10% and 13%, respectively, while identical applications from Russian and English backgrounds have 23% and 27% response rates. Thus, we focus on immigrants from Africa or Western Asia.[4]

[3] These numbers come from Statistics Finland's public StatFin database.

[4] We do not include Russians because the long and complex history between Finland and likely means attitudes towards them originate from different sources than those towards Africans and Asians. Specifically, Finland was part of the Russian Empire in 1809–1917, fought against the Soviet Union during World War II, and endured heavy influence from the Soviet Union during the Cold War. In addition, a large fraction of immigrants born in Russia or the Soviet Union are ethnic Finns (Ingrians), who were granted a return migrant status after the collapse of the Soviet Union and are much more positively regarded than ethnic Russians (Jaakkola, 2000, 2005) but who we cannot distinguish from ethnic Russians in our data.

3. DATA AND MEASUREMENT

3.A. Data sources

Statistics Finland constructed our data by linking together information from several administrative registers. The most important, and novel, information is the precise residential location of every person living in Finland at the end of each year between 1988 and 2017. The resolution of this location information depends on local population density. We observe building-level co-ordinates for everyone living in a building with more than a certain number of residents. For smaller housing units, we observe residential location primarily at the level of 250×250m grid cell.[5] Most of the persons included in our baseline sample, described below, live in densely populated areas and thus 62% of individuals are assigned to a building, and 38% are assigned to a 250m x 250m grid cell.

Our data also contain annual information on individual characteristics, educational attainment, employment status, and income for everyone living in Finland. In addition, we observe identification codes for individuals' households and family units. Importantly, these data allow us to identify children of immigrants based on parents' place of birth regardless of whether the child is born in Finland or abroad.

3.B. Baseline sample

We focus on non-immigrant individuals born between 1977 and 1999. We exclude older birth cohorts from the analysis because we measure exposure to immigrant neighbors between ages 5 and 15 using data that starts in 1987. Similarly, we exclude younger cohorts, because our outcome variables measure behavior in adulthood, and our data ends in the year 2017. That is, we observe the youngest birth cohorts until age 18.

In addition, we exclude individuals growing up in neighborhoods where the share of immigrants is below 1 %.[6] This restriction increases statistical power since it excludes a larger number of individuals who are very unlikely to meet potential immigrant partners as adults, and also helps to deal with issues arising from the distribution of the immigrant share being essentially censored at zero. Overall, this leaves us with a sample of 236,515 individuals.

[5] In the relatively rare case that fewer than 40 individuals live in the 250×250m grid cell, we observe a 1×1 kilometer grid cell instead.

[6] More precisely, we restrict based on the share of individuals living in a neighborhood between the ages of five and fifteen who were born within five years of the person who could potentially be in our sample. We show that our results are not sensitive to this exact cutoff in Table **??**.

3.C. Treatment variable

Our aim is to understand how childhood exposure to peers of immigrant background affects later outcomes. We do not directly observe childhood interactions and thus need to proxy likely interactions using information available in our data. We therefore exploit the fact that children tend to socialize with their neighbors, partly through going to the same school, and are most likely to associate with those of the same age and gender (Kalmijn, 2002; Merlino et al., 2019). Thus, we use the immigrant share among children of the same age and gender growing up in the same neighborhood as our exposure measure.

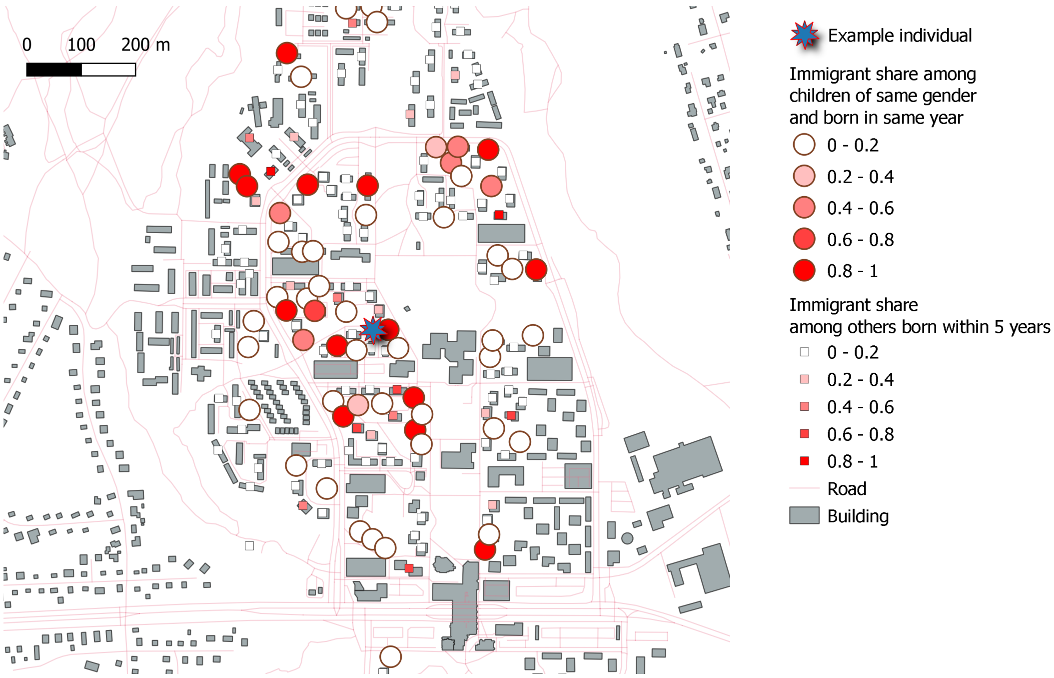
To operationalize this strategy, we first have to define a 'neighborhood'. This process is not straightforward—we want to capture as closely as possible the group of children that a given child is most likely to interact with. It would be simple, for instance, to define a neighborhood as a grid cell of some arbitrary size, but this has two major problems. First, a grid cell of a given size will vary substantially in the number of children who live within it depending on the density of the area. Second, defining neighborhoods in this way will introduce substantial noise since those living on the edge of a given cell would be more likely to interact with children in an adjacent cell than those living on the far side of their own cell. We therefore construct a unique neighborhood for every location on which we have information. A location's neighborhood is then defined as the closest set of locations that contain one hundred people per birth cohort.[7] Neighborhoods are also restricted not to contain locations further than 5km away, even if this means they will contain fewer than one hundred people on average, but this restriction is generally not binding. An example of a neighborhood is given in Figure 1.

For each neighborhood, we then calculate the immigrant share for each birth cohort for each year and average across the years for which that cohort was between 5 and 15 years old to give the average childhood immigrant share for each cohort in each neighborhood. We then assign each person to an individual 'childhood location', which is the modal location in which they lived at ages 5–15. In our analysis, we do not use information on other locations an individual may have lived in—although this may well be relevant for influencing their behavior, assigning individuals to multiple childhood locations would make the analysis more complex and less transparent.[8] From Table 1, we can see that, on average, 3% of an individual in our sample's childhood neighbors are immigrants.

[7] Table 1 shows that the average person in our sample has around 44 people of the same gender and 43 people of the opposite gender in their individual cohort. The difference stems from the fact we are including the individual themselves.

[8] The median person in our sample spent 70% of the years between ages 5 and 15 in their modal childhood location.

FIGURE 1: EXAMPLE OF A COHORT IN A NEIGHBORHOOD



Notes: To preserve anonymity, this figure doesn't display an actual cohort but instead generates a synthetic cohort based on the real data for a collection of 11 cohorts. In particular, for each location, we calculate the average number of immigrants and non-immigrants per cohort, and then generate a random cohort by drawing from Poisson distributions with these means.

3.D. Outcome variables

Our main outcome variable is defined as whether an individual ever lived with an immigrant partner.[9] From Table 1, we can note that around 1% of individuals in our sample have lived with an immigrant partner, though it is important to note that over 40% have not lived with any partner, which reflects the fact that many of the individuals in our sample are fairly young when we last observe this outcome variable.

We also analyze the impact of individuals' childhood cohort immigrant share on the immigrant share among their location when they are adults, as well as the share of their colleagues and managers. To construct the first variable, for each year

[9]We define partners through marriage and cohabitation. Statistics Finland defines cohabitation as two unmarried adults of different sex aged 18 and over and occupying the same dwelling on a permanent basis, provided their age difference is less than 16 years and they are not siblings. In case the couple has a common child these specifications do not apply. Persons of the same sex living together are not regarded as cohabiting couples unless the household-dwelling unit consists of two women who are the biological mother of the child living in the same dwelling and the second mother confirmed for the child.

TABLE 1: SUMMARY STATISTICS

Variable	Mean	Std. Dev.
Birth year	1992.054	4.877
Male	0.491	0.5
Number of people of same gender in neighbourhood cohort	43.726	10.41
Number of people of opposite gender in neighbourhood cohort	42.853	10.393
Same gender immigrant share in neighbourhood cohort	0.031	0.036
Opposite gender immigrant share in neighbourhood cohort	0.032	0.038
Immigrant share in neighborhood within 5 years	0.032	0.029
Ever lived with a partner	0.565	0.496
Ever lived with an immigrant partner	0.01	0.101
N	236515	

Notes: An individual's neighbourhood cohort is made up of the children born in the same year living closest to the location the individual lived between the ages of 5 and 15 - see Section 3.C for more details.

after the person left their childhood home, we look at the immigrant share within the 250m grid cell containing the individual's place of residence.[10]

4. EMPIRICAL STRATEGY

Our identification strategy exploits variation across cohorts within locations in the immigrant share of children living in the neighborhood. For instance, suppose that within a neighborhood there are 100 children in each birth cohort, and that on average two percent are immigrants. Since we are dealing with relatively small numbers, there will be important relative variation - it is likely, for instance, that at least one cohort has only one immigrant, while another will have three. By comparing non-immigrants in the cohort with those in the cohort with three, we can observe whether quasi-random variation in the cohort immigrant share impacts non-immigrants' outcomes later in life.

All of our analyses consist of regressing outcomes of interest on the neighborhood immigrant share in the relevant birth cohort while controlling for location and cohort-gender fixed effects. The key identifying assumption is that, within a given location, variation in immigrant share across birth cohorts is uncorrelated with other variables that might influence our outcomes of interest. This assumption might not hold, however, if variation is driven by changes over time in the general immigrant share within a location. In most locations, variation driven by such changes is likely to be much smaller than the natural quasi-random variation across cohorts. Nonetheless, in order to deal with this concern, we use three alternative specifications that include extra components to address this issue.

[10]The variable is therefor undefined for those who have not moved out of their childhood home, but we find no impact of one's childhood immigrant share on whether and individual has moved out.

We begin by undertaking a graphical analysis where we also include the immigrant share of individual cohorts up to five years younger or older than the individual's cohort. Plotting the coefficients on each of these cohorts allows us to observe whether that on the individuals particular cohort is likely to be driven by confounding trends. In other words, we plot the coefficients β_t and γ_t of the following regression:

$$Y_i = \sum_{t=-5}^{5} \beta_t ImmShare^s_{l,g,c+t} + \sum_{t=-5}^{5} \gamma_t ImmShare^o_{l,g,c+t} + I_l + I_{cg} + \varepsilon_i \tag{1}$$

where Y_i is the outcome of individual i, $ImmShare^s_{l,g,c}$ is the share of immigrants among the same-gendered birth cohort c who grew up in location l, $ImmShare^s_{l,g,c}$ is the share of immigrants among the opposite-gendered birth cohort c who grew up in location l, I_l is a set of location fixed effects, and I_{cg} a set of birth cohort by gender fixed effects .

For our main tables, we then use two alternative specifications. First, we include as an additional control a flexible function of the immigrant share of the children in birth cohorts up to five years away from an individuals' cohort. In particular, we split the variable up into ten linear splines. Since this is the immigrant share amongst a much larger group (around one thousand people) it should control for any other confounding variables correlated with the cohort immigrant share. Second, we include location-specific trends in addition to location fixed effects. This is arguably more flexible than the first approach, but also reduces the amount of variation we can use since in many locations our sample is fairly small. Consistent with this approach, we show in Table A1 in the Appendix that the variation we use with either of these specification isn't systematically correlated with a set of predetermined variables.

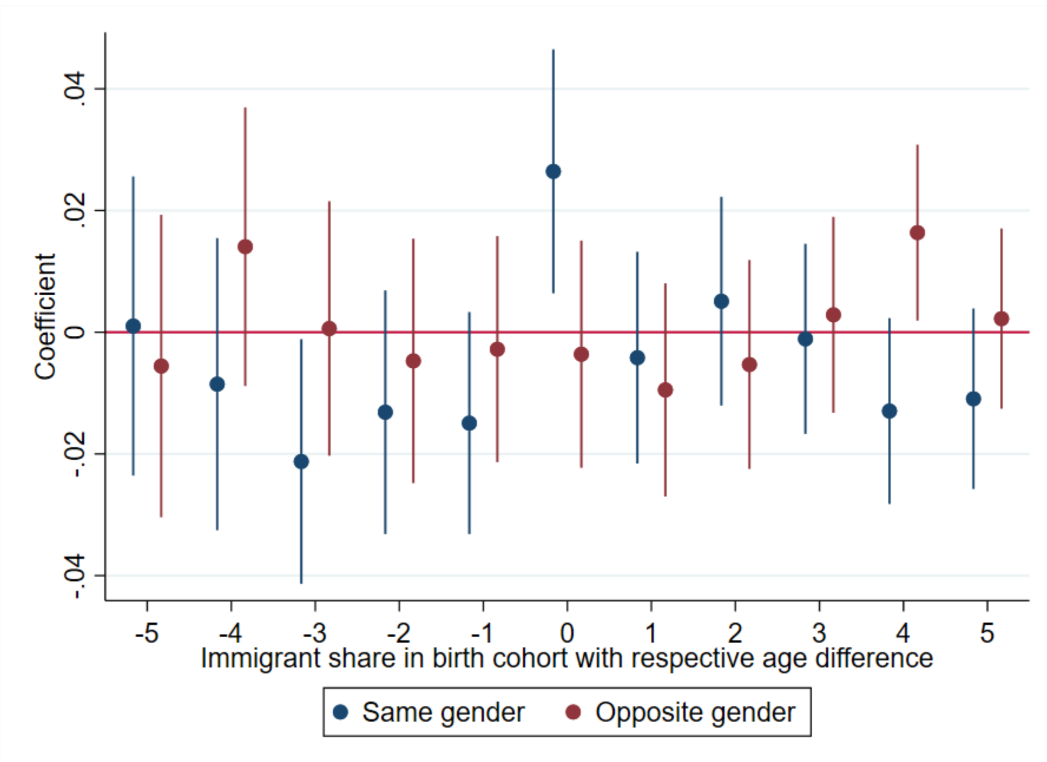
For our main tables, we show results both combining and splitting the cohort neighborhood immigrant share by gender. For some outcomes, it may be ex ante unclear which share will have a larger impact. The opposite gender share may be more relevant if this is a pool of potential partners, for instance, but the same gender share may be more relevant if these people are more likely to become friends. It is very difficult to think, however, of a confounding variable that is likely to be correlated with the cohort immigrant share among one gender but not the other. We can therefore be fairly confident that any consistent difference between the coefficients on the same and opposite gender shares results from an impact of these shares on individuals' behavior.

5. Results

5.A. *Main results*

We begin with a graphical analysis based on plotting the estimated coefficients of equation 1 - results are displayed in Figure 2. From the figure, it is immediately noticeable that one coefficient has a larger magnitude than the others, and is significantly different from zero - namely, the coefficient on the cohort of the same age and same gender. Moreover, there do not appear to be any trends among the other coefficients, backing up our identification assumption that, once we control for location and cohort fixed effects, any correlation between cohort immigrant share and our outcome of interest is likely to be causal.

FIGURE 2: RELATIONSHIP BETWEEN CHILDHOOD COHORT IMMIGRANT SHARE AND ADULT COHABITATION, BY AGE AND GENDER DIFFERENCE



The magnitude of the estimated coefficient is substantial. Increasing the immigrant share in an individual's cohort by one standard deviation - about 3.6 percentage points - increases the likelihood of them having an immigrant partner by 0.1 percentage points - i.e. 10% of the sample mean. This is very close to the magni-

Table 2: Impacts of cohort immigrant share on probability of cohabiting with an immigrant partner

	Full sample		Men		Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Cohort immigrant share	0.0248	0.0274	0.0234	0.0206	0.00938	0.0275
	(0.0166)	(0.0207)	(0.0166)	(0.0243)	(0.0335)	(0.0457)

Cohort immigrant share - same gender	0.0281***	0.0304**	0.0175*	0.0180	0.0356*	0.0274
	(0.0106)	(0.0132)	(0.01000)	(0.0135)	(0.0214)	(0.0308)
Cohort immigrant share - opposite gender	-0.00114	-0.00120	0.00618	0.00464	-0.0212	0.00214
	(0.00983)	(0.0131)	(0.0120)	(0.0189)	(0.0197)	(0.0288)
f(11-year imm share)	Y		Y		Y	
Location specific trends		Y		Y		Y
Observations	236515	236515	112136	112136	107423	107423
Locations	34842	34842	23399	23399	22729	22729
Dep. var. mean	0.0104	0.0104	0.00382	0.00382	0.0175	0.0175

Notes: The dependent variable takes the value one if an individual has cohabited with an immigrant partner, and zero otherwise. A cohort is a group of children born in the same year as the individual living in a nearby location between the ages of 5 and 15 - see Section 3.C for more details. Childhood location and cohort-gender fixed effects are included in all columns. In columns 1, 3, and 6 we additionally include ten linear splines of the average immigrant share among children in the same area born within 5 years of the individual. In columns 2, 4, and 6 we include childhood location trends. Standard errors are clustered by childhood location. * $p < .10$, ** $p < .05$, *** $p < .01$

tude found in Merlino et al. (2019) who look at similar variation where the ethnic minority in question is blacks in the US.

Table 2 then presents the results from our two alternative specifications. Odd-numbered columns control for a flexible function of the immigrant share among children living in the neighborhood born within five years of the individual, while even-numbered columns control for location-specific trends. In the first two columns we look at the results in the whole sample, as in the graphical analysis, and we find very similar results. In the remaining columns, we split the sample by gender to observe whether the effect may be driven by men or women. Results indicate the impact is significant for both genders, and differences between the coefficients across columns are insignificant.

5.B. Robustness

We undertake a number of tests to measure the robustness of our results. In our main analysis, we restricted our sample to individuals living in locations where the share of immigrants born within five years was greater than one percent, and we check to ensure that the results don't vary greatly if we change this threshold. In Figure A1, we show the results of our analysis for a range of other thresholds, and find that our results don't change substantially. While the coefficient is lower when we include areas with lower immigrant shares, it is also the case that the mean of the dependent variable is lower, presumably at least partly because people in these areas are much less likely to meet potential immigrant partners.

Since our choice of neighbourhood size was also somewhat arbitrary, we also test how sensitive our results are to changes in neighbourhood size in Figure A2. Results remain very similar if we increase the neighbourhood size, but become smaller and insignificant if we halve the size of a neighbourhood to an average size. This suggests that the larger neighbourhood is likely to be the more relevant group when it comes to the exposure that we are concerned with, which is beleivable given the small geographical size of even the relatively larger neighbourhoods.

We may also be concerned that our treatment variable is clearly somewhat correlated across space - if an individual in one location is exposed to a high immigrant share, others in the same cohort living nearby are also likely to be. Although we wouldn't necessarily expect similar spatial correlation in our outcome variable, we test the sensitivity of our results to alternative clustering techniques in Table A3. We can note that our coefficient of interest remains significant whether we cluster at the 1km grid cell level or use Conley spatial standard errors with a range of different distance thresholds.

Finally, we check to see whether our coefficients move substantially when we introduce control variables - notably those that we used to check for balance in Table A1. In Table A4 we can note that including these variables barely moves our coefficients, suggesting that our results are unlikely to be driven by omitted variable bias.

5.C. Other outcomes and mechanisms

Given the significant impact of childhoood exposure on intermarriage, a natural question that arises is whether there are impacts on other aspects of an individual's life. In particular, are they more likely to have immigrants as neighbors, colleagues, or managers? Table 3 suggest this is not the case - coefficients on the other outcomes we examine are small and insignificant.[11] One potential explanation for this difference in impacts is that individual prejudices may play less of a role in decisions over location or workplace, at least on the margin of prejudice that may be impacted by childhood contact. In any case, these results suggest that an increased opportunity to meet immigrants as an adult through one's residential location or workplace is not driving the results.

An alternative potential mechanism would be that an increase in immigrants in one's childhood neighborhood has an impact on educational performance, thereby

[11] We find similar results looking at other alternative measures of the same outcomes, such as the immigrant share within an individual's building or 1km grid cell.

TABLE 3: IMPACTS OF COHORT IMMIGRANT SHARE ON OTHER OUTCOMES

	Immigrant share of 250m grid cell		Immigrant share of colleagues		Immigrant share of managers	
	(1)	(2)	(3)	(4)	(5)	(6)
Cohort immigrant share	0.00310	0.00600	0.0117	0.0143	0.0198	0.0250
	(0.0125)	(0.0161)	(0.0116)	(0.0149)	(0.0209)	(0.0262)
Cohort immigrant share - same gender	0.00279	0.00767	0.00285	0.00193	0.00590	0.0124
	(0.00874)	(0.0111)	(0.00782)	(0.00975)	(0.0149)	(0.0187)
Cohort immigrant share - opposite gender	-0.000541	-0.00233	0.00777	0.0137	0.0139	0.0154
	(0.00717)	(0.0102)	(0.00657)	(0.00950)	(0.0123)	(0.0168)
f(11-year imm share)	Y		Y		Y	
Location specific trends		Y		Y		Y
Observations	152373	152373	150996	150996	148714	148714
Locations	26646	26646	26496	26496	26259	26259
Dep. var. mean	0.0419	0.0419	0.0172	0.0172	0.0126	0.0126

Notes: A cohort is a group of children born in the same year as the individual living in a nearby location between the ages of 5 and 15 - see Section 3.C for more details. Childhood location and cohort-gender fixed effects are included in all columns. In columns 1, 3, and 6 we additionally include ten linear splines of the average immigrant share among children in the same area born within 5 years of the individual. In columns 2, 4, and 6 we include childhood location trends. Standard errors are clustered by childhood location. * $p < .10$, ** $p < .05$, *** $p < .01$ * $p < .10$, ** $p < .05$, *** $p < .01$

influencing the share of immigrants an individual meets during their education or changing their opportunities on the dating market. We therefore examined whether there is any impact of exposure on educational outcomes including an individual's grade average, whether they ever attended tertiary education, and whether or not they were in employment, education, or training in 2017. We find no significant impacts on any of these variables (see Table A2 for more details).

It is therefore likely that the results we find on immigrant partners stems from at least one of two potential mechanisms. First, it may be that individuals with a greater childhood immigrant exposure are more likely to make friends with immigrants as children, and as a result have a greater share of immigrants among their social networks which persist into adulthood. This higher share of immigrants among individuals' social networks then increases their probability of meeting an immigrant who they then go on to have a romantic relationship with. Second, exposure to immigrants of the same gender at school may change individuals' attitudes towards immigrants and therefore change their probability of becoming friends or partners with an immigrant as an adult even outside of their childhood friendship network.

Since we do not observe individuals' social networks or attitudes, we cannot estimate which of these mechanisms is likely to be the most important. Nonetheless, we undertake a couple of exercises which shed some light as to whether the effect we find is likely to be entirely driven by direct social network effects.

First, for both middle and high schools, we split an individuals childhood neighborhood cohort into those who went to the same school and those who went to a different school. Doing so, we find that there are significant impacts of both the same-school and other-school groups (Table A5 in the appendix presents these results). This suggests that, if the effect is being driven by social networks formed as children, an important part of these networks must be being made and maintained outside of school.

Second, in Table 4 we restrict the pool of relationships that we consider in the dependent variable to those which we believe are less likely to be directly the result of social networks formed in school. In columns 1 and 2, we include only relationships with immigrants that begun at least two years after the individual had moved more than 50km away from their childhood location. In columns 3 and 4, we include only relationships that begun with someone who had never lived within 50km of the individual's childhood location. In both cases, of course, it is not impossible that the individual met their partner through a social network formed in the childhood location, but this mechanism is unlikely. Importantly, we still see significant impacts of childhood immigrant exposure on both of these outcomes, which suggests that to some extent there is likely to have been a change in individuals' preferences or attitudes.

6. Conclusion

This article has demonstrated that exposure to immigrants as children increases individuals' propensity to cohabit with immigrant partners as adults. Evidence suggests that such an effect is at least partly due to a change in individuals' preferences or attitudes, consistent with the contact hypothesis. We can therefore reasonably conclude that reducing residential segregation may aid the integration of minority groups partly by changing the way the majority group behaves towards them.

Of course, while the intermarriage rate between a minority group and the majority are an important sign of integration, it is not typically the direct aim of policy makers. It is reasonable to believe, however, that the change of social networks or attitudes that greater intermarriage implies may also lead to further integration in terms of a broader set of political, social, and economic outcomes. While in this paper we did not find any impact on other outcomes, this is likely to be at least in part because of a lack of data and relative scarcity of immigrants in the Finnish context. Future work can therefore shed further light on the extent to which childhood contact impacts other long-term outcomes besides intermarriage.

TABLE 4: IMPACT ON RELATIONSHIPS WITH IMMIGRANTS FORMED FURTHER AWAY

	Began after moved away		Spouse never lived nearby	
	(1)	(2)	(3)	(4)
Cohort immigrant share	0.0109***	0.00933**	0.0126**	0.0121*
	(0.00381)	(0.00468)	(0.00546)	(0.00671)
Cohort immigrant share - same gender	0.00954***	0.00804**	0.0121***	0.0104**
	(0.00261)	(0.00332)	(0.00347)	(0.00435)
Cohort immigrant share - opposite gender	0.00123	0.00148	0.000864	0.00223
	(0.00228)	(0.00259)	(0.00298)	(0.00365)
f(11-year imm share)	Y		Y	
Location specific trends		Y		Y
Observations	236515	236515	236515	236515
Locations	34842	34842	34842	34842
Dep. var. mean	0.000786	0.000786	0.00129	0.00129

Notes: The dependent variable in columns 1 and 2 takes the value one if an individual has cohabited with an immigrant partner and this cohabitation begun at least two years after the individual moved more than 50km away from their childhood location. The dependent variable in columns 3 and 4 takes the value one if an individual has cohabited with an immigrant partner who has never lived within 50km of the individual's childhood location. Both dependent variables are zero if the individual has not had the relevant kind of relationship. A cohort is a group of children born in the same year as the individual living in a nearby location between the ages of 5 and 15 - see Section 3.C for more details. Childhood location and cohort-gender fixed effects are included in all columns. In columns 1 and 3 we additionally include ten linear splines of the average immigrant share among children in the same area born within 5 years of the individual. In columns 2 and 4 we include childhood location trends. Standard errors are clustered by childhood location. * $p < .10$, ** $p < .05$, *** $p < .01$

REFERENCES

Ahmad, Akhlaq, "Do Equal Qualifications Yield Equal Rewards for Immigrants in the Labour Market?," *Work, Employment and Society*, 2020, *34* (5), 826–843.

Allport, Gordon W, "The Nature of Prejudice," *New York: Addison*, 1954.

Bazzi, Samuel, Arya Gaduh, Alexander D Rothenberg, and Maisy Wong, "Unity in diversity? How intergroup contact can foster nation building," *American Economic Review*, 2019, *109* (11), 3978–4025.

Billings, Stephen B, Eric Chyn, and Kareem Haggag, "The Long-Run Effects of School Racial Diversity on Political Identity," *American Economic Review: Insights*, 2021.

Brown, Jacob R., Ryan D. Enos, James Feigenbaum, and Soumyajit Mazumder, "Childhood cross-ethnic exposure predicts political behavior seven decades later: Evidence from linked administrative data," *Science Advances*, 2021, *7* (24), eabe8432.

Bursztyn, Leonardo, Thomas Chaney, Tarek Alexander Hassan, and Aakaash Rao, "The Immigrant Next Door: Exposure, Prejudice, and Altruism," Technical Report, National Bureau of Economic Research 2021.

Dustmann, Christian, Kristine Vasiljeva, and Anna Piil Damm, "Refugee migration and electoral outcomes," *The Review of Economic Studies*, 2019, *86* (5), 2035–2091.

Furtado, Delia and Stephen J Trejo, "Interethnic marriages and their economic effects," in "International handbook on the economics of migration," Edward Elgar Publishing, 2013.

Holmlund, Helena, Erica Lindahl, and Sara Roman, "Immigrant peers in the class: responses among natives and the effects on long-run revealed preferences," 2021.

Jaakkola, Magdalena, "Finnish attitudes towards immigrants in 1987-1999," *Finnish Yearbook of Population Research*, 2000, pp. 129–161.

— , *Suomalaisten suhtautuminen maahanmuuttajiin vuosina 1987-2003*, Ministry of Labour, Helsinki, 2005.

Kalmijn, Matthijs, "Intermarriage and homogamy: Causes, patterns, trends," *Annual review of sociology*, 1998, *24* (1), 395–421.

— , "Sex Segregation of Friendship Networks. Individual and Structural Determinants of Having Cross-Sex Friends," *European Sociological Review*, 2002, *18* (1), 101–117.

Meng, Xin and Robert G Gregory, "Intermarriage and the economic assimilation of immigrants," *Journal of Labor economics*, 2005, *23* (1), 135–174.

Merlino, Luca Paolo, Max Friedrich Steinhardt, and Liam Wren-Lewis, "More than just friends? School peers and adult interracial relationships," *Journal of Labor Economics*, 2019, *37* (3), 663–713.

Paolini, Stefania, Fiona A White, Linda R Tropp, Rhiannon N Turner, Elizabeth Page-Gould, Fiona K Barlow, and Ángel Gómez, "Intergroup contact research in the 21st century: Lessons learned and forward progress if we remain open," *Journal of Social Issues*, 2021, *77* (1), 11–37.

Schindler, David and Mark Westcott, "Shocking racial attitudes: black GIs in Europe," *The Review of Economic Studies*, 2021, *88* (1), 489–520.

Steinmayr, Andreas, "Contact versus Exposure: Refugee Presence and Voting for the Far Right," *Review of Economics and Statistics*, 2021, *103* (2), 310–327.

A. Appendix: Extra tables

Table A1: Balance table

	Parents age		Parents have secondary education		Parents have tertiary education		Imm share of parents' colleagues		Parental income (pctile)		At least one parent in data	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Cohort immigrant share	-0.0875	-0.867	0.0171	0.0520	-0.118**	-0.108*	-0.00862	-0.00739	-6.315*	-1.177	0.0166	0.0919***
	(0.870)	(1.100)	(0.0527)	(0.0674)	(0.0463)	(0.0582)	(0.00621)	(0.00786)	(3.666)	(4.630)	(0.0276)	(0.0343)
Cohort immigrant share - same gender	0.422	0.0382	0.0124	0.0322	-0.0185	-0.0154	-0.00496	-0.00251	-1.507	1.198	0.0157	0.0539**
	(0.585)	(0.725)	(0.0357)	(0.0449)	(0.0316)	(0.0393)	(0.00400)	(0.00456)	(2.445)	(3.041)	(0.0189)	(0.0229)
Cohort immigrant share - opposite gender	-0.354	-0.786	0.00816	0.0241	-0.0874***	-0.0848**	-0.00422	-0.00485	-4.433*	-2.339	0.0113	0.0373*
	(0.554)	(0.720)	(0.0334)	(0.0446)	(0.0288)	(0.0377)	(0.00428)	(0.00598)	(2.330)	(3.056)	(0.0174)	(0.0213)
f(11-year imm share)	Y		Y		Y		Y		Y		Y	
Location specific trends		Y		Y		Y		Y		Y		Y
Observations	187472	187472	187472	187472	187472	187472	148715	148715	187472	187472	236515	236515
Locations	31617	31617	31617	31617	31617	31617	27652	27652	31617	31617	34842	34842
Dep. var. mean	30.73	30.73	0.772	0.772	0.182	0.182	0.00233	0.00233	46.40	46.40	0.802	0.802

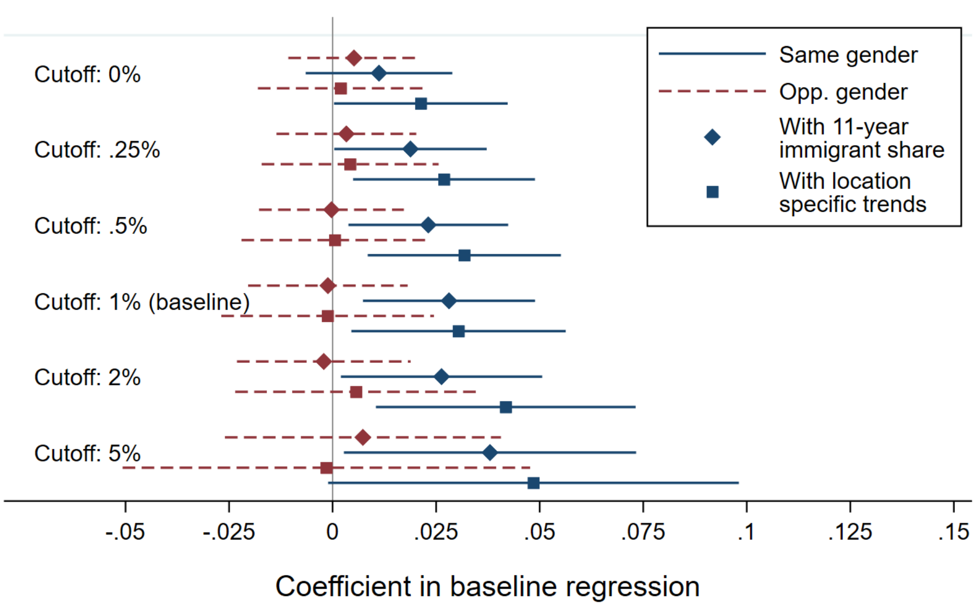
Notes: A cohort is a group of children born in the same year as the individual living in a nearby location between the ages of 5 and 15 - see Section 3.C for more details. Childhood location and cohort-gender fixed effects are included in all columns. In odd numbered columns we additionally include ten linear splines of the average immigrant share among children in the same area born within 5 years of the individual. In even numbered columns we include childhood location trends. Standard errors are clustered by childhood location. * $p < .10$, ** $p < .05$, *** $p < .01$

TABLE A2: IMPACTS OF COHORT IMMIGRANT SHARE ON EDUCATIONAL OUTCOMES

	Middle school grade point average		Not in education, employment, or training		Ever attended tertiary education	
	(1)	(2)	(3)	(4)	(5)	(6)
Cohort immigrant share	-0.0581	-0.0229	0.0828	0.0872	0.0266	0.180
	(0.164)	(0.196)	(0.0660)	(0.0829)	(0.114)	(0.156)
Cohort immigrant share - same gender	0.00103	0.0589	0.0533	0.0310	0.0276	0.162
	(0.112)	(0.134)	(0.0446)	(0.0561)	(0.0830)	(0.108)
Cohort immigrant share - opposite gender	-0.106	-0.0419	0.0348	0.0560	-0.0246	0.0300
	(0.0992)	(0.123)	(0.0392)	(0.0530)	(0.0640)	(0.0923)
f(11-year imm share)	Y		Y		Y	
Location specific trends		Y		Y		Y
Observations	217709	217709	173259	173259	127367	127367
Locations	33326	33326	27979	27979	22301	22301
Dep. var. mean	7.710	7.710	0.127	0.127	0.471	0.471

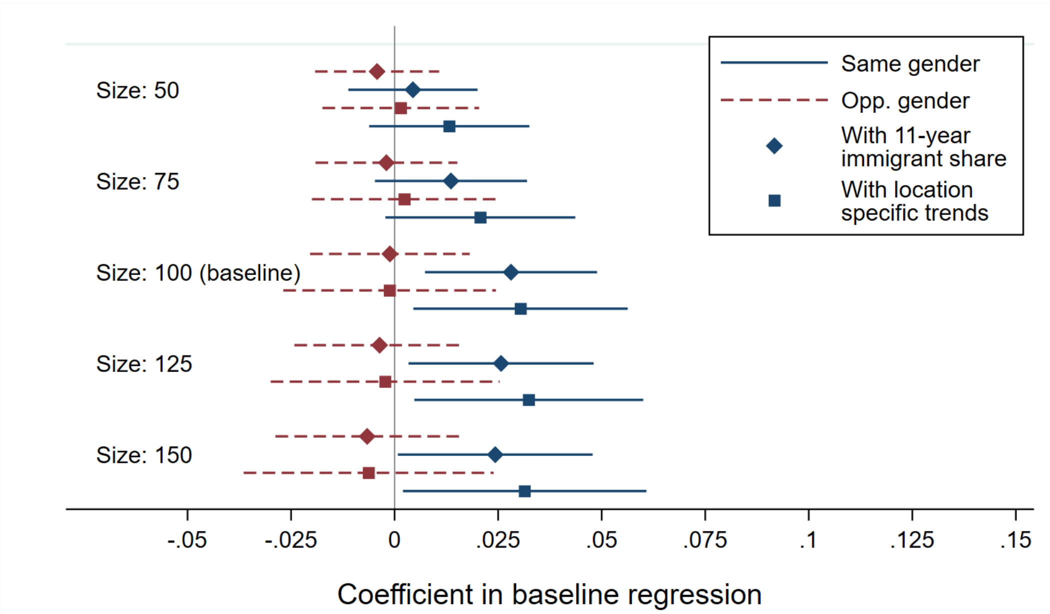
Notes: A cohort is a group of children born in the same year as the individual living in a nearby location between the ages of 5 and 15 - see Section 3.C for more details. Childhood location and cohort-gender fixed effects are included in all columns. In columns 1, 3, and 6 we additionally include ten linear splines of the average immigrant share among children in the same area born within 5 years of the individual. In columns 2, 4, and 6 we include childhood location trends. Standard errors are clustered by childhood location. * $p < .10$, ** $p < .05$, *** $p < .01$

FIGURE A1: ROBUSTNESS TO VARYING THE SAMPLE SELECTION CUTOFF



Notes: This figure displays the coefficients in the baseline regression - i.e. the second panel of columns 1 and 2 in Table 2 - when we use alternative thresholds on the immigrant share among the group born within 5 years of the individual.

FIGURE A2: ROBUSTNESS TO VARYING THE NEIGHBORHOOD SIZE



Notes: This figure displays the coefficients in the baseline regression - i.e. the second panel of columns 1 and 2 in Table 2 - when we use alternative neighborhood sizes.

Table A3: Robustness to alternative clustering

	Cluster at 1km grid cell		Conley spatial standard errors		
			3km	5km	10km
	(1)	(2)	(3)	(4)	(5)
Cohort immigrant share	0.0248	0.0274	0.0248*	0.0248*	0.0248*
	(0.0162)	(0.0195)	(0.0146)	(0.0146)	(0.0146)
Cohort immigrant share - same gender	0.0281***	0.0304**	0.0281***	0.0281***	0.0281***
	(0.0106)	(0.0118)	(0.00956)	(0.00934)	(0.00919)
Cohort immigrant share - opposite gender	-0.00114	-0.00120	-0.00114	-0.00114	-0.00114
	(0.00994)	(0.0129)	(0.00897)	(0.00890)	(0.00880)
f(11-year imm share)	Y		Y	Y	Y
Location specific trends		Y			
Observations	236515	236515	236515	236515	236515
Locations	34842	34842	34842	34842	34842
Dep. var. mean	0.0104	0.0104	0.0104	0.0104	0.0104

Notes: The dependent variable takes the value one if an individual has cohabited with an immigrant partner, and zero otherwise. A cohort is a group of children born in the same year as the individual living in a nearby location between the ages of 5 and 15 - see Section 3.C for more details. The size of the neighbourhood corresponds to the average number of children per cohort - larger neighbourhoods include locations further from an individual's childhood location. Childhood location and cohort-gender fixed effects are included in all columns. In columns 1 and 3-5 we additionally include ten linear splines of the average immigrant share among children in the same area born within 5 years of the individual. In column 2 we include childhood location trends. Standard errors are clustered by the 1km grid cell in columns 1 and 2, and using Conley spatial standard errors in columns 3-5. * $p < .10$, ** $p < .05$, *** $p < .01$

Table A4: Robustness to including control variables

	(1)	(2)
Cohort immigrant share	0.0249	0.0279
	(0.0166)	(0.0207)
Cohort immigrant share - same gender	0.0284***	0.0309**
	(0.0106)	(0.0132)
Cohort immigrant share - opposite gender	-0.00122	-0.00110
	(0.00983)	(0.0131)
f(11-year imm share)	Y	
Location specific trends		Y
Observations	236515	236515
Locations	34842	34842
Dep. var. mean	0.0104	0.0104

Notes: The dependent variable takes the value one if an individual has cohabited with an immigrant partner, and zero otherwise. A cohort is a group of children born in the same year as the individual living in a nearby location between the ages of 5 and 15 - see Section 3.C for more details. The size of the neighbourhood corresponds to the average number of children per cohort - larger neighbourhoods include locations further from an individual's childhood location. Childhood location and cohort-gender fixed effects are included in all columns. In column 1 awe additionally include ten linear splines of the average immigrant share among children in the same area born within 5 years of the individual. In column 2 we include childhood location trends. In both columns we additionally control for the variables used as dependent varialbes in Table A1. Standard errors are clustered by childhood location. * $p < .10$, ** $p < .05$, *** $p < .01$

TABLE A5: SPLITTING BY SCHOOL

Cohort immigrant share				
- same school	0.0849***	0.0984***	0.000270	0.0100
	(0.0294)	(0.0358)	(0.0469)	(0.0637)
Cohort immigrant share				
- diff school	0.000417	-0.00693	0.0290*	0.0301
	(0.0117)	(0.0140)	(0.0158)	(0.0205)
Cohort immigrant share	0.0661***	0.0725**	0.0534	0.0507
- same gender, same school	(0.0240)	(0.0298)	(0.0340)	(0.0465)
Cohort immigrant share	0.0216*	0.0254*	0.0244*	0.0316*
- same gender, diff school	(0.0112)	(0.0140)	(0.0125)	(0.0165)
Cohort immigrant share	0.0270	0.0373	-0.0487	-0.0360
- opposite gender, same school	(0.0237)	(0.0313)	(0.0360)	(0.0467)
Cohort immigrant share	-0.00984	-0.0179	0.00646	0.00555
- opposite gender, diff school	(0.0101)	(0.0123)	(0.0122)	(0.0159)
f(11-year imm share)	Y		Y	
Location specific trends		Y		Y
Observations	236209	236209	166626	166626
Locations	34819	34819	29656	29656
Dep. var. mean	0.0104	0.0104	0.00910	0.00910

Notes: The dependent variable takes the value one if an individual has cohabited with an immigrant partner, and zero otherwise. A cohort is a group of children born in the same year as the individual living in a nearby location between the ages of 5 and 15 - see Section 3.C for more details. Childhood location and cohort-gender fixed effects are included in all columns. In columns 1, 3, and 6 we additionally include ten linear splines of the average immigrant share among children in the same area born within 5 years of the individual. In columns 2, 4, and 6 we include childhood location trends. Standard errors are clustered by childhood location. * $p < .10$, ** $p < .05$, *** $p < .01$