The Stockholm University
Linnaeus Center for
Integration Studies (SULCIS)

*Peers, neighborhoods and immigrant student achievement – evidence from a placement policy**

Olof Åslund, Per-Anders Edin, Peter Fredriksson & Hans Grönqvist

Working Paper 2011:2
ISSN 1654-1189

Peers, neighborhoods and immigrant student achievement - evidence from a placement policy[*]

By

Olof Åslund[#], Per-Anders Edin[⊗], Peter Fredriksson[♣], and Hans Grönqvist[⊕]

November 2010

Abstract

We examine to what extent immigrant school performance is affected by the characteristics of the neighborhoods that they grow up in. We address this issue using a refugee placement policy which provides exogenous variation in the initial place of residence in Sweden. The main result is that school performance is increasing in the number of highly educated adults sharing the subject's ethnicity. A standard deviation increase in the fraction of high-educated in the assigned neighborhood raises compulsory school GPA by 0.9 percentile ranks. Particularly for disadvantaged groups, there are also long-run effects on educational attainment.

Keywords: Peer effects; Ethnic enclaves; Immigration; School performance
JEL-codes: J15; I20; Z13

[*] We are grateful to Thomas Lemieux (the editor) and two anonymous referees, as well as Miles Corak, David Cutler, Richard Freeman, Andrea Ichino, Per Johansson, Kevin Lang, Mikael Lindahl, Daniele Paserman, Nicole Schneeweis and Eskil Wadensjö for helpful comments and suggestions. We have also benefited from comments by seminar and conference participants at The Institute for Fiscal Studies, University of Padova, Harvard University, University of Kalmar, Uppsala University, Stockholm University, IEB (Barcelona), the Nordic Summer institute in Labor Economics (Aarhus), the Nordic Migration Workshop (Helsinki), and the Workshop on the Economics of Immigration (Vancouver). Special thanks to Louise Kennerberg and Jörgen Moen for preparing the data.
[#] Institute for Labour Market Policy Evaluation (IFAU), Uppsala University, IZA, and Uppsala Center for Labor Studies (UCLS). Address: IFAU, P.O. Box 513, SE-751 20 Uppsala, Sweden. Email: olof.aslund@ifau.uu.se.
[⊗] Uppsala University, IFAU, and UCLS. Address: Department of Economics, Uppsala University, P.O. Box 513, SE-751 20 Uppsala, Sweden. Email: per-anders.edin@nek.uu.se.
[♣] Stockholm University, IZA, and UCLS. Address: Department of Economics, Stockholm University, SE-106 91 Stockholm, Sweden. Email: peter.fredriksson@ifau.uu.se.
[⊕] Swedish Institute for Social Research (SOFI), Stockholm University. Address: SOFI, SE-106 91 Stockholm, Sweden. Email: hans.gronqvist@sofi.su.se.

1 Introduction

Immigrant students typically perform substantially worse than native students in the OECD countries. According to PISA, the performance gap between first generation immigrants and natives amounts to around half a standard deviation in math, reading, and science (OECD 2006a). In this paper we examine to what extent this is due to the characteristics of the neighborhoods in which the immigrants grow up. Since recently arrived immigrants tend to settle in close proximity to people sharing their ethnic background (Stark 1991), we pay particular attention to the characteristics of the ethnic community.

There is a large literature on the impact of residential and school segregation on the outcomes of disadvantaged groups in general. But there is not so much dealing with immigrant children in particular. This is perhaps surprising given that the characteristics of the neighborhood community can exert particularly strong influences on young migrants striving to find their place in the new country. Moreover, the work by Heckman and coauthors (e.g., Cunha and Heckman 2007) suggests that the impact of the environment is more pronounced in disadvantaged families.

The question we examine also sheds light on the rationale for policies designed to shift the location of immigrants. These policies may come in the form of incentive programs, such as Moving to Opportunity (see Kling et al. 2007), or deliberate attempts by the governments to restrict the location choices of new immigrants; the latter kind of policies are (or have been) practiced by many European countries (see Edin et al. 2004).

It is an open question whether the characteristics of the ethnic community has a causal effect on immigrant student achievement. Ethnic concentration per se may be

beneficial if the enclave provides useful information on, e.g., the workings of the education system, but detrimental if residential concentration hampers proficiency in the host country's language. But the characteristics of the contacts are arguably at least as important. Well-established and educated peers may act as role models, but living among people with poor socioeconomic status and performance may have a negative influence on youth (cf. Cutler and Glaeser 1997).

Our paper is related to several strands of literature. First, there is a large literature on the impact of residential segregation on adult minorities (including immigrants) in general.[1] The evidence is somewhat mixed. Segregation per se may hurt individuals (e.g. Cutler and Glaeser 1997) but the literature also points to the importance of the quality of neighborhood contacts (Bertrand et al. 2000; Åslund and Fredriksson 2009).

Second, there is a growing body of (largely U.S.) research studying the effects of racial composition within schools or neighborhoods on student performance.[2] In general, these studies suggest that the performance of black students is reduced by attending schools with a large fraction of black students.

Third, there is a small literature examining whether ethnic concentration affects the school performance of immigrants. Cortes (2006) studied the effect of age at arrival and attending an enclave school on the test scores of a sample of first and second generation immigrants residing in the cities of Miami and San Diego in the U.S. She found that

[1] See, e.g., Åslund and Fredriksson (2008), Beaman (2009), Bertrand et al. (2000), Cutler and Glaeser (1997), Edin et al. (2003), and Goel and Lang (2009) for recent contributions.
[2] See, e.g., Angrist and Lang (2004), Boozer et al. (1992), Card and Rothstein (2007), Grogger (1996), Guryan (2004), Hanushek et al. (2009), Hoxby (2000), and Rivkin (2000).

attending an enclave school (defined as one where above 25 percent are foreign-born) had no effect on students' test scores.[3]

Fourth, there are some studies which examine whether immigrants' labor market success is related to the characteristics of the childhood neighborhood.[4] For instance, Borjas (1995) found that (second generation) immigrants who grew up in ethnic communities with an abundance of human capital did better on the labor market.

The studies by Cortes (2006) and Borjas (1995) are directly relevant to our paper. However, as for many other studies of contextual effects, one could worry that selection problems bias the estimates in these two studies. This is mainly because a student's neighborhood or school is a family choice variable. If residential choice is based on unobserved characteristics which also affect learning outcomes, the estimates will be biased and cannot be interpreted causally.

We rely on a governmental placement policy to generate exogenous variation in the initial residential distribution. During 1987–91 Swedish authorities assigned refugees to their initial location. Since individuals were not free to choose, we argue that the initial location was independent of (unobserved) individual characteristics, an issue we will obviously return to below.[5]

Our strategy is demanding on data availability. We have access to administrative records containing detailed information on all students graduating from Swedish com-

[3] See Bygren and Szulkin (2007) for a related study using Swedish data. Jensen and Rasmussen (2008) have examined whether student outcomes are related to immigrant concentration using Danish data. Their estimates suggest a negative impact of immigrant concentration on student performance. Neither of these studies in practice handles the problems caused by residential self-selection.

[4] The paper by Grönqvist (2006) also belongs to this category.

[5] We have previously used this approach to study economic outcomes among adult migrants; see Edin et al. (2003) Åslund and Fredriksson (2009) Åslund et al. (2006) and Åslund and Rooth (2007). Gould et al. (2004) use a similar placement policy where Ethiopian refugees were distributed across Israeli municipalities to identify the causal effect of school quality on students' high school grades. There are also papers exploiting similar policies in Denmark; see e.g. Damm (2009a, 2009b).

pulsory schools during 1988–2003, and we are able to track their educational success beyond compulsory school. The data also contain rich individual information on the population aged 16–65 (starting in 1985), and provide the opportunity to link children to their parents. This means that we can identify when the individual arrived, where he or she initially resided, the characteristics of his or her parents, and the properties of the neighborhood peers at different points in time.

A first set of results suggests that a larger ethnic community has a positive impact on school performance, whereas the size of the overall immigrant population in the neighborhood is negatively related to compulsory school GPA. However, the latter estimate is not identified with neighborhood fixed effects and is potentially subject to omitted variables bias.

In our main analysis we therefore focus on the impact of the size and characteristics of the ethnic community. The results suggest that a standard deviation increase in the fraction of highly educated peers in the assigned neighborhood raises compulsory school GPA by 0.9 percentile ranks; a corresponding increase in the size of the ethnic community in the assigned neighborhood has about the same effect, but the effect is less precisely estimated. The effects of the characteristics of the ethnic community are larger among those who arrived before age seven than for those who arrive at an older age.

Had we not accounted for residential self-selection using the placement policy, our conclusions regarding the impact of ethnic concentration would have been very different. Auxiliary regressions suggest that disadvantaged children (in the unobserved sense) are sorted into neighborhoods with a high share of members from their own ethnic group. The sorting bias is so severe that the size of the ethnic community at the time of

graduation is negatively related to student outcomes. Sorting bias does not plague the estimate on the educational composition of the ethnic group, however.

The analysis also shows that the effects of the educational composition of peers do not vary by gender or parental education. However, the size of the ethnic community is more important for boys and for children whose parents are less-educated, two groups that have the poorest school outcomes. Moreover, for these two groups we find that the characteristics of the assigned location influence long-run educational attainment: A large and highly educated community means a significantly higher probability of graduating from upper secondary education at the normal age, which is a strong predictor of obtaining a university degree.

The remainder of this paper is organized as follows. The next section provides background information on the educational system, how immigrant students perform in Swedish schools, and the placement policy which we base our analysis on. In Section 3, we present the data. Section 4 outlines the empirical strategy in more detail and contains the empirical results. Section 5 concludes.

2 Background

2.1 Immigration and residential concentration in Sweden

Sweden has a large immigrant population: 12 percent (out of a population of 9 million) are foreign-born. Even though Sweden has received net migration since the 1930s, the larger inflows began in the 1950s and 1960s as workers were recruited primarily from Finland, but also from Central and Southern Europe and Turkey. Starting in the 1970s, labor migrants were gradually replaced by refugees and family reunification migrants, a development which accelerated in the 1980s and 1990s. The large refugee inflows have

changed the source country composition of the immigrant population dramatically. Parallel to these demographic changes there has been a decline in the economic performance of migrants. Today, Sweden stands out as one of the countries with the largest immigrant-native differentials in the labor market (OECD 2007).

As in other Western countries, the immigrant population is concentrated to certain regions and neighborhoods. Greater Stockholm, Göteborg and Malmö host about one third of the overall population but as much as half of the foreign-born. Within larger regions, immigrants tend to be concentrated to particular areas, usually situated in the suburbs (Åslund et al. 2010). The residential concentration is also reflected in the immigrant share of the neighborhoods populated by the foreign-born.[6] The typical immigrant lives in an area where a quarter of the working-age population is foreign-born, which can be compared to the national average of 12 percent.

Previous studies show that the typical immigrant-dense neighborhood contains a mix of ethnic groups. Such areas are primarily united by a shortage of natives (Andersson 2000). Still, different groups are relatively concentrated in different areas; e.g. Iranians constitute a substantially larger share of the foreign-born in Göteborg than in Sweden's other major cities. Also at the finest geographic level this segregation is evident; people have substantially more country-of-origin peers living in their neighborhood than what can be explained by regional sorting or by a division of immigrants and natives in general. We return to this issue in the description of our sample of child migrants.

[6] As described in the data section we use SAMS (Small Area Market Statistics) areas to define neighborhoods.

2.2 Immigrants in Swedish schools

Compulsory education is 9 years in Sweden and starts at age 7; the typical age at graduation is thus 16.[7] There is a national curriculum that all compulsory schools follow. After compulsory school a vast majority go on to upper-secondary education where admission is based on compulsory school grades. Completing three years of upper-secondary education is a prerequisite for enrolling at the universities.

We study cohorts graduating from the nine-year compulsory school between 1988 and 2003. Within this time-frame, the grading system was reformed. Until 1998, grades given at graduation were relative, with a fixed national average for each graduating cohort. From 1998, grades are on an "absolute" scale, which is to be based on performance only and not related to the achievement of others. Because of this grading reform we transform the data to percentile ranks of the individual grade within cohort.[8]

Of special interest for our study are the rules for allocating students to schools. Up until 1991, the Swedish compulsory school system assigned students to the school situated nearest to their residential area. This residence principle is still the leading rule on how to allocate students to schools. However, in 1992, the central government introduced a school choice reform, where parents in principle are free to choose their children's school within the municipality. It is important to note, however, that parental preferences are severely constrained by space limitations, and priority is given to kids residing close to the school. Thus, the assignment of refugee children to neighborhoods to a very large degree determined which schools they attended. Also, since there are far

[7] See Björklund et al. (2005) for further details on the Swedish education system.
[8] Transforming the data to percentile ranks also deals with grade inflation in the new system (e.g., Cliffordson, 2004).

more neighborhoods than schools, controlling for area of residence effectively also means controlling for schools.

There is ample evidence that immigrant children perform poorly in the Swedish school system. According to PISA 2003, the gap between the Swedish-born and the foreign-born at age 15 amounts 0.7–0.8 standard deviations of the PISA score distribution in math, reading and science (OECD 2006a). The gap between the native-born and immigrants is about twice as large as the gender difference in reading. Within the immigrant group, there are big differences depending on time spent in Sweden: those who arrive after age 7 perform substantially worse than those who migrate before age 7 (Böhlmark 2008).

2.3 The refugee placement policy[9]

In 1985, the Swedish Immigration Board was given the task of assigning newly arrived refugee immigrants to an initial municipality of residence. The policy was introduced in response to complaints from cities that had experienced a rise in immigration and perceived this as a burden on local public budgets. By placing asylum seekers in municipalities that had suitable characteristics for reception the government hoped to speed up the integration process.

Because of the large inflow of asylum seekers in the late 1980s, the number of receiving municipalities was increased from 60 to include 277 of Sweden's 284 municipalities in 1989. Available public housing essentially determined the placement. The policy was formally running 1985–94, but the implementation was strictest between

[9] Edin et al. (2003) contains a more detailed description of the placement policy. Their description of the practical implementation of the placement policy is based on interviews with placement officers and other officials at the Immigration Board.

1987 and 1991. During this period, the placement rate was around 90 percent, and the individuals involved were given very little room to choose the initial municipality of residence. Therefore, we focus our analysis on the period 1987–91.

Asylum seekers were placed in refugee centers pending a decision from the immigration authorities. The centers were located all over Sweden, and center assignment was independent of port of entry to Sweden. The mean duration between entry into Sweden and the receipt of a permit varied between three and twelve months during 1987–91. After receiving the permit, municipal placement occurred within a much shorter period of time, partly because there were explicit goals for reducing the time span between receipt of the residence permit and placement. Refugee preferences were considered in the municipal assignment, but individuals applied for residence in the largest cities where there were few vacancies because of the economic boom. Assigning a refugee to a municipality was conditional on having found a vacant apartment within that particular municipality. (Since individuals were assigned to an apartment, they were in practice assigned to a neighborhood.) After having been assigned to an apartment, refugees were basically free to move. The only "cost" of moving, apart from direct moving costs, was delayed enrolment in language courses.

2.3.1 Placement as a policy experiment

The *a priori* arguments for considering placement as exogenous with respect to the unobserved characteristics of the individual are the following: (i) the individual could not choose his or her first place of residence due the institutional setup, the practical limitations imposed by scarce housing, and the short time frame between the receipt of residence permit and placement; (ii) there was no direct interaction between local place-

ment officers and individual refugees, meaning that any selection must have been on observed characteristics.

With respect to the first point, note that the timing of the receipt of the residence permit must coincide with the arrival of a housing vacancy in the preferred location, if the refugee was to realize his or her most preferred option. The joint probability of these two events happening at the same time is extremely low.[10] Thus immigrant preferences are likely to have played a very limited role in the assignment process.

The above argument does not guarantee that immigrants were randomly assigned to neighborhoods. In situations where there are more immigrants receiving a residence permit than there are housing vacancies, it is still possible that there is selection by municipal officers. In such cases, the selection was purely in terms of observed characteristics. Interviews with officials involved in the assignment process indicate that language, education, and family size were the crucial characteristics. Preferences were given to highly educated individuals and individuals who spoke the same language as some members of the resident immigrant stock; single individuals were particularly difficult to place, since small apartments were extremely scarce. It is important to note that we have information on all these crucial characteristics in our data.

The geographic distribution of immigrant before and after introduction of the placement policy substantiates the argument that the policy generated a location distribution which was independent of unobserved individual characteristics. Edin et al. (2003) showed that the overall geographic distribution of those subjected to the placement policy differed substantially from the location choices made by migrants

[10] Oreopoulos (2003) uses a similar argument to motivate why assignment to a public housing project can be considered exogenous for new recipients of welfare payments in Toronto.

arriving from the same regions shortly before the reform. Åslund et al. (2010) showed that the initial characteristics of the assigned locations differed pre and post reform; but after 9–10 years in Sweden the sorting pattern of those who arrived under the placement policy came to resemble that of other migrants. We take this as evidence that people were not able to realize their preferred option.

To test of our key identifying assumption we have regressed initial neighborhood characteristics on pre-determined parental and child characteristics. Since parental education, family size, and country of origin were potentially used in the placement process, one should not be surprised if one would find that some of these characteristics are correlated with the characteristics of the neighborhood. However, we have found no indications that pre-determined characteristics of the child were used in the placement process. Therefore, we base our test of the assumption that placement was exogenous conditional on the observables on the estimate on child age at immigration. Note that migration age has an independent and quantitatively important effect on school performance (see Böhlmark 2008; Bleakley and Chin 2004).

Figure 1 shows the relationship between age at migration and the share of highly educated in the ethnic community ("ethnic human capital") in the assigned location. Ethnic human capital and age at migration are completely unrelated. A regression of (the log of) the share of high-educated on age at migration entered linearly produces an estimate of 0.04 with a standard error of 0.08; an analogous regression with the (log) size of the ethnic community as the dependent variable yields an estimate of 0.004 (standard error: 0.02) on age at migration.[11]

[11] In a previous version of the paper we demonstrated that the birth month of the child was also unrelated to initial neighborhood characteristics; see Åslund et al (2009).

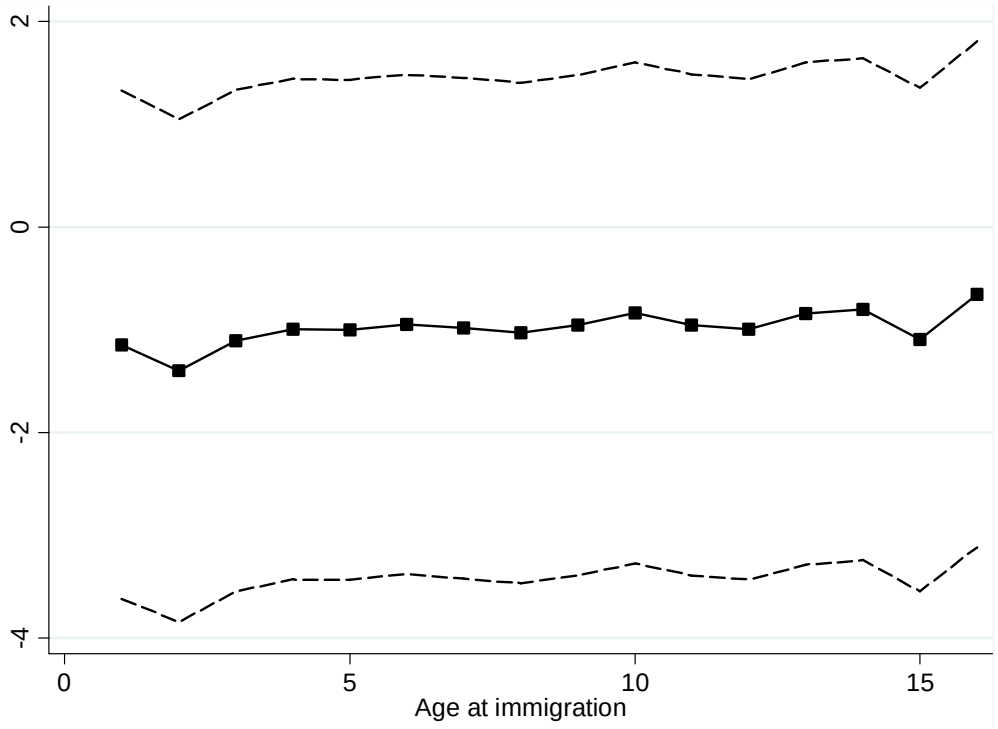

Figure 1 Ethnic human capital in assigned location by age at immigration

Notes: The figure shows estimates in log points (solid line, 95 percent confidence interval given by dashed lines) from a linear regression of the log share of highly educated in the ethnic community in the assigned location on a set of age at immigration dummies. The model also controls for gender, age of the mother, the educational attainment of the mother and the father, as well as fixed effects for family size, country of birth, neighborhood immigration year, and graduation year.

For completeness, Table A1 in the appendix reports the estimated correlations between initial neighborhood characteristics and pre-determined parental and child characteristics (see column (1) and (3)). None of the included parental and child characteristics are significantly related to the share of high-educated at the time of arrival. Three (out of 11) characteristics are significantly related to the size of the community. All in all, we think that the results reported in Table A1 lend additional

support to the argument that the placement policy generated exogenous variation in neighborhood assignment.[12]

Given the institutional setting, and the information documented here, we think it is valid to assume that the assignment location is exogenous to the child, conditional on family background and family size. Note that this assumption is less strict than in, e.g., Edin et al. (2003), since child and parental characteristics are not perfectly correlated.[13]

3 Data

We use administrative data covering the entire Swedish population aged 16–65. The data originate from administrative registers maintained by Statistics Sweden and contain information on, e.g., labor market status, educational attainment, income, taxes, and various demographic variables.[14] An important feature of the data is that we can link students to their parents and we are thereby able to include information on several parental characteristics. We define parental characteristics separately for each parent.

Our sample consists of the refugee children who graduated from compulsory school 1988–2003 and whose parents obtained their residence permit 1987–91.[15] From 1988 and onwards there is information on all final grades for students graduating from

[12] In the regressions pertaining to initial neighborhood characteristics 3 out of 22 coefficients are significant at the 10 percent level. We think this supports our case since just by chance 2 out of 22 coefficients would be significant at the 10 percent level. Columns (2) and (4) in Table A1 reports the results of analogous exercise, this time pertaining to the time of graduation. These estimates are clearly inflicted by sorting bias, as illustrated by the fact that 14 of the 22 regressions coefficients are significant.

[13] Estimates of the intergenerational earnings correlation are typically much lower in Sweden than in the U.S. Corak (2006) reports estimates for different countries: the estimate for Sweden is 0.27 compared to 0.47 for the U.S.

[14] The key registers are the income tax registers (*Inkomst- och taxeringsstatistiken*), population registers (*Registret för totalbefolkningen*), the register on educational attainment (*Utbildningsregistret*), the grade 9 register (*Årskurs-9 registret*), and the multi-generational register (*Flergenerationsregistret*).

[15] The sample was created by first identifying individuals from the relevant source countries who graduated from compulsory school 1988–2003 who immigrated (i.e. obtained their residence permit) 1987–91. Then we added information on all their family members. An individual was retained in the sample if the child entered Sweden at the same time point as the parent. We condition on there being an identified mother in the family. This sampling procedure creates the complete set of parent-child pairs who entered the country in the same year sometime during 1987–91, conditional on there being an identified mother in the family. The identity and characteristics of the father is sometimes missing as explained in the main text.

Swedish compulsory school. The individuals were between 0 and 16 years of age at immigration. We identify refugee immigrants by region of origin and exclude children who did not arrive together with the parent who first came to Sweden. The motivation for excluding these individuals is that they are likely to have immigrated for family reunification reasons, and these immigrants were exempted from the placement policy.

In this paper we use SAMS (Small Area Market Statistics) areas to capture neighborhoods. SAMS areas are defined as homogenous areas in certain respects; it may be a homogenous area with certain types of buildings—high-rise buildings, owner-occupied housing, or business complexes, for instance. SAMS areas are the smallest geographic units available in Swedish data. Sweden has about 9,000 SAMS areas, which gives an average of 1,000 residents (of which about 600 are aged 16–65). However, the average individual lives in an area with 1,849 inhabitants aged 16–65. Since the foreign-born are concentrated to urban areas it is not surprising to find that the average immigrant lives in a somewhat more populated area; the average immigrant lived in a SAMS area with 2,498 inhabitants aged 16–65.

Since individuals do not enter the data before age 16, we use the assignment location of the parent(s) who arrived together with the child to get information on the first SAMS area. We also measure the characteristics of the location observed in the individual's year of graduation. A potential problem is that we only observe the region of residence at the end of the year. If the observed initial location differs from the actual initial placement due to internal migration, this creates a measurement error in initial placement. This issue has been thoroughly investigated in Edin et al. (2003) where a weighting scheme based on aggregate data on municipal refugee reception from the Immigra-

tion Board was used. The estimates from the weighted regressions were very similar to the non-weighted ones, suggesting that this measurement error is not a big concern.

Notice that, by and large, schools aggregate neighborhoods. There are close to 2,000 schools and 9,000 SAMS areas. In principle, it would be interesting to examine whether it is the characteristics of the neighborhood or the school which matter for student achievement. But in practice it will be very hard to disentangle the two. Since the characteristics of the neighborhood will capture the neighborhood as well as the schools, we choose to measure the characteristics at the neighborhood level.[16]

We study two educational outcomes. The first and primary one is the percentile rank (by graduation year) of the compulsory school GPA. Although not perfect, the GPA is the best widely available summary measure of compulsory school performance in Sweden. Furthermore, it is the basis for admission and selection to upper secondary school. The second outcome is the probability of completing upper secondary school ″on time″ (i.e. by age 19 which is the normal graduation age). Ideally we would have liked to examine university enrollment or graduation rates. But a substantial fraction of our sample is too young for such an analysis to be meaningful. Notice, however, that having a degree from upper-secondary school is a prerequisite for enrolling at the university and that graduating on time is a strong predictor of future success in the education system.[17]

[16] There is some scope for trying to disentangle the effects of school and neighborhood characteristics. Children in some neighborhoods go to different schools, and there is time variation in school catchment areas. But given that there are substantial difficulties in identifying catchment areas, we leave this endeavor for future research.

[17] To substantiate the latter claim we have examined how the probability of having a university degree by age 29 is related to the probability of graduating on time. In the overall population of individuals born in 1976 with a degree from upper-secondary school, the probability of having at least a 4-year degree from university is 58 percent higher if the individual graduated on time (18.8% relative to 11.9%); the probability of having at least a 2-year university degree is 36 percent higher for those who graduated on time relative to those who graduated later than normal (56.4% relative to 41.6%).

3.1 A description of the sample

Table A2 and Table A3 provide some general descriptive statistics of the estimation sample, containing a total of 20,039 individuals. Not unexpectedly, outcomes are quite poor: the average percentile rank of GPA is 40 and only 43 percent graduate from upper-secondary school on time[18]. The typical child migrant in the sample was 8 years old at migration. There are slightly more boys in the sample (53–47) and mean sibship size is close to 3, which is relatively high by Swedish standards.

A fair share (16.5 percent) of the fathers is not present in the data. Among those observed, educational information is unavailable for about 11 (7.6) percent of the fathers (mothers). The observed distribution of education shows that about half the parents have only compulsory education. Thirty percent have some short or long upper-secondary education, and approximately 20 percent have obtained education at the university level.

It is also clear that there is variation in region of origin. Iranians are the largest group, contributing a quarter of the sample. 17.8 percent originate in Northern Africa and 13.3 percent in Chile. About 20 percent of the individuals have arrived from different parts of Eastern Europe and the former USSR.

The descriptive statistics also show residential concentration among the studied refugees. There is substantial variation in the size of the SAMS population in the sample, but the average is higher than what is observed in the overall population, which is consistent with concentration to larger cities with higher population density. The immigrant share in the neighborhood (at the time of graduation) is as high as 31 percent, which is much higher than in the overall population. Concentration in the "ethnic" dimension is

[18] The equivalent number is 67 percent among Swedish-born individuals in the 1981 cohort (which corresponds to the average birth year of our sample of refugee immigrants).

even stronger: on (a weighted) average, the groups studied constitute 0.6 percent of the working-age population, yet the average "ethnic" share in the neighborhood is 3.2 percent at the time of graduation.

4 How do neighborhood characteristics affect immigrant student achievement?

We begin this section by discussing specification issues and our empirical strategy. We pursue two different specifications. One is designed to estimate the impact of the size of the immigrant community, the other to estimate the impact of the characteristics of the ethnic community, holding immigrant concentration constant. The latter specification constitutes our main empirical approach.

We then turn to presenting the results. Section 4.2 examines the impact of the size of the immigrant community in the assigned location. Section 4.3 presents the results pertaining to the characteristics of the ethnic community; the section contains the average effects as well as separate estimates by certain observed characteristics (gender, parental education, and age at arrival). Sections 4.2 and 4.3 focus on a reduced-form approach where we relate initial neighborhood characteristics to later student outcomes. The advantage of this approach is that it estimates a well-defined causal effect while imposing a minimal set of assumptions. A disadvantage of the reduced-form approach is that it is harder to pinpoint why the characteristics the initial neighborhood is of importance. In section 4.4 we therefore impose additional structure by assuming that it is average exposure to neighborhood characteristics, between the time of arrival and the time when we measure outcomes, which is of importance for educational outcomes. Since average exposure is endogenous, we use the characteristics of the initial location

as instruments. This IV approach particularly facilitates the interpretation of the estimates for various sub-groups since it corrects for differential mobility rates across groups. The drawback, of course, is that we have to assume that the characteristics of the initial assignment are excludable conditional on average exposure.

In the presentation of the results we focus mainly on percentile ranked GPA. But in section 4.3 we also examine whether neighborhood characteristics matter for the probability of graduating from upper-secondary school on time.

4.1 Empirical strategy and specification issues

To fix ideas, consider the following simple model (where we have suppressed arrival time fixed effects and graduation time fixed effects for convenience).

$$y_{ics} = \alpha x_i + \beta^e \ln X_{cs}^e + \beta^m \ln X_s^m + \beta^p \ln X_s^p + \lambda_s + \lambda_c + \varepsilon_{ics} \qquad (1)$$

where *i* indexes individuals, *c* countries of origin, and *s* neighborhoods (SAMS areas). *y* is the outcome of interest, X^j, $j = e, m, p$, denotes the characteristics of the (*e*)thnic community, the (*m*)igrant community, and the (*p*)opulation in the neighborhood. x_i denotes a vector of individual characteristics (the subject's age at immigration, the mother's age, mother's and father's level of education, gender and family size).

Notice that the effects of X_{cs}^e are identified even if we treat λ_s as neighborhood fixed effects, since there is variation across ethnicities within a neighborhood. However, the effects of X_s^m and X_s^p are not, since there is no variation within a neighborhood. This obvious point demonstrates a trade-off in the analysis: investigation of some issues comes at the price of stronger assumptions for identification.

Indeed, a lot of the (European) policy discussion focuses on the consequences of attending immigrant dense schools or growing up in immigrant dense neighborhoods. To tackle this wider policy question, we replace the neighborhood fixed effects with municipality fixed effects (there are 290 municipalities). The effects of X_s^m, say, are then identified using the variation across neighborhoods within a municipality. The estimates from this specification will not suffer from bias due to individual self-selection, given that the placement policy generates variation in neighborhood characteristics which are independent of unobserved individual characteristics. But there is a potential for bias due omitted variables at the neighborhood level, for instance, due to correlations between unobserved school quality and immigrant density.[19]

The neighborhood fixed effects model imposes a weaker set of assumptions. Therefore we focus on this model and thus elaborate mostly on the importance of the characteristics of the ethnic community.

4.2 The impact of the size of the immigrant and ethnic communities

Table 1 reports the results of a barebones model, where we relate immigrant student achievement to the sizes of the ethnic and immigrant communities. In column (1) we present the results from the municipality fixed effects model, while column (2) contains the neighborhood fixed effects model. Throughout we enter the neighborhood characteristics in logs.[20]

[19] In principle, there is also a risk that the model with neighborhood fixed effects suffers from omitted variables bias. But since the omitted variable would have to vary across ethnicity within neighborhood, we do not think this is a big concern. Moreover, the direction of any bias is unclear.

[20] The log specification is very convenient since it implies that the results are invariant to the precise segregation measure used; see Bertrand et al. (2000) on this point. Although convenient, the log specification comes with a small "price". We encounter some problems when there are no fellow countrymen in the community. We deal with this issue by assigning an arbitrary low value for the size of the ethnic community and then include a dummy variable that

Table 1 A barebones model

	Dependent variable: Percentile ranked GPA	
	(1)	(2)
Characteristics measured at year of arrival		
Size of ethnic community	.646**	.514*
	(.247)	(.290)
Size of immigrant community	–1.034**	
	(.524)	
Population size	.879	
	(.554)	
(Initial) SAMS FE:s	No	Yes
(Initial) Municipality FE:s	Yes	No
Ethnic group FE:s	Yes	Yes
Year of arrival FE:s	Yes	Yes
Year of graduation FE:s	Yes	Yes
Number of observations	20,039	20,039

Notes: Neighborhood characteristics are measured in logs. The sample consists of refugee immigrants whose parents arrived during the period 1987−1991 and who completed compulsory school not later than 2003. All regressions control linearly for the subject's and the mother's age, with dummies for each parent's educational attainment (five levels), family size, gender and missing values. Standard errors robust for clustering at the SAMS*ethnic group level (5947 cells) in parentheses. ** = significant at 5 % level; * = significant at 10 % level

The estimates in column (1) suggest a positive impact of a larger ethnic community. By contrast, there is a negative effect of expanding the immigrant community. Notice that the estimate on the size of the ethnic community captures the effect of replacing an immigrant of another ethnicity with an immigrant of the subject's own ethnicity (since the overall size of the immigrant community is held constant). The estimate on the size of the immigrant community, on the other hand, should be interpreted as the effect of increasing the density of immigrants of another ethnicity (since the size of the ethnic community is held constant).

The interpretation of the estimates in column (1) relies on the assumption that we have not omitted relevant neighborhood variables. Including neighborhood fixed effects in column (2) only marginally changes the coefficient on the size of the ethnic

indicates no other fellow countrymen. Note that the inclusion of the dummy variable implies that the procedure of assigning arbitrary values to empty cells will not affect the estimate on the neighborhood characteristics. Further, the estimate on the size of the community gives the effect of increasing the size of the community conditional on there being at least one person from one's own ethnic group in the neighborhood.

community, which can be taken to suggest that omitted variables are not such a big concern.

How should the magnitudes be interpreted? Since the neighborhood variables are entered in logs, a unit change corresponds to increasing the size of the community by around 170 percent.[21] Evaluated at this change, an increase in the size of the ethnic community in the assigned location has the effect of raising immigrant student achievement (at graduation) by 0.65 percentile ranks. An increase in the density of other immigrants would reduce immigrant performance by roughly a percentile rank. On the basis of the estimates, we can also examine what happens to student performance when the size of the ethnic group changes, taking into account that this will also change overall immigrant density. The effect of increasing the size of the ethnic community, holding only neighborhood population constant, equals 0.56 which is significant at the 5-percent level (the standard error is 0.23).

4.3 The impact of the characteristics of the ethnic community

Now, let us turn to the impact of the characteristics of the ethnic community. To analyze this issue we focus on the specification including neighborhood fixed effects, a specification which is robust to omitted variables at the neighborhood level.

Column (2) of Table 1 reports the estimates of a model including only the size of the own community. As noted above, the effect of increasing the size of the ethnic community in the assigned location is positive. But the result does not yield so much insight into why this is the case.

[21] This is just to say that (exp(1)–exp(0)) ≈ 1.7. Notice that the standard deviation of the log of the size of the ethnic group is 1.3, i.e., it exceeds unity.

To make some headway into the question of why the neighborhood matters we postulate what we think of as a pure peer effects model. Our incarnation of this model is that the student outcomes of immigrant children are influenced by the educational background of the children with whom they potentially interact, in school as well as in the neighborhood. In practice we assume that X_{cs}^{e} = (the number of highly educated adults with kids under age 18 in the ethnic community).[22] It is straightforward to decompose this quantity into three components: (i) the number of adult countrymen (aged 25–65) living in the neighborhood (denoted by *N*); (ii) the fraction of these countrymen who are high-educated, i.e. have at least three years of upper-secondary education (which is denoted by *h*); and (iii) the fraction of the highly-educated countrymen in the neighborhood who have kids under age 18 (denoted π). We thus have $X_{cs}^{e} = (N \times h \times \pi)_{cs}^{e}$. Introducing this expression into equation (1), and attaching a separate coefficient on the components, we get

$$y_{ics} = \alpha x_i + \beta_1^e \ln N_{cs}^e + \beta_2^e \ln h_{cs}^e + \beta_3^e \ln \pi_{cs}^e + \lambda_s + \lambda_c + \varepsilon_{ics} \qquad (2)$$

where we have suppressed X_s^m and X_s^p since they do not vary within neighborhood and are thus picked up by the fixed effects. We emphasize again that the neighborhood variables are measured at the time of immigration, since this is the only time when neigborhood characteristics are exogenous to the unobserved characteristics of the individual. Moreover, we exclude the parent(s) of the individual when calculating the neighborhood characteristics.

[22] We would have liked to have a closer matching between the age of the subject (the immigrant child) and the age range of his potential peers. Since the ethnic communities are so small this not feasible in practice.

The specification in (2) provides a convenient test of what characteristics of the ethnic community are important, and to some extent why. If $\beta_1^e = \beta_2^e = \beta_3^e$, the pure peer effects model applies and it is the number of highly educated parents that have an impact on student performance. The configuration $\beta_3^e = 0, \beta_1^e = \beta_2^e$ may suggest that the neighborhood is important because all adults act as role models. In this case, it is the number of highly educated in the entire ethnic community that matters; there is no additional effect coming from the human capital of the parents. In general, β_2^e measures the impact of increasing the human capital of the community while holding size constant, while β_1^e gives the effect of increasing the size of the community while holding the educational composition constant.

This specification can be seen as a way of estimating the impact of the assignment location invoking a minimum of assumptions. An alternative view is to interpret equation (2) as a reduced form of a structural model where school performance is affected by cumulated peer influences between the time of immigration and the time of graduation (see section 4.4 and Åslund and Fredriksson 2009).

4.3.1 Baseline results

Table 2 presents the baseline results relating compulsory school GPA to neighborhood characteristics. The table only reports the results of main interest; the estimates on the other included characteristics are presented in Table A4. These additional covariates exhibit the expected impact. Girls outperform boys by about 8 percentile ranks on average. Parental education has a substantial impact on outcomes: children with university educated mothers outperform children who have mothers with compulsory education by

11 points (the estimates on father's education have a similar flavor). There are substantial performance differences across birth regions and also patterns suggestive of worse outcomes in larger families, even though these patterns are weaker than what is sometimes found in descriptive studies (Åslund and Grönqvist 2010).

Let us now turn to the estimates of the upper panel of Table 2, where school performance is related to the characteristics of the assigned neighborhood. Both the size and the educational attainment of the ethnic community have a positive impact on performance. There is no additional effect coming from the human capital of the parents. The latter result may be somewhat surprising. One interpretation is that highly educated adults in the ethnic community act as role models.

Table 2 The relationship between neighborhood characteristics and compulsory school grades

	Dependent variable: Percentile ranked GPA		
	(1)	(2)	(3)
Panel A. Year of arrival			
Size of ethnic community	.647*	.488	.409
	(.330)	(.310)	(.315)
Share with high education	1.141**	.987**	1.120**
	(.511)	(.498)	(.508)
Share of high-educated who are parents	–.209	--	--
	(.668)		
Interaction (size and share high-educated)	--	--	–.078
			(.059)
Panel B. Year of graduation			
Size of ethnic community	–.522**	–.532**	–.680**
	(.228)	(.196)	(.207)
Share with high education	1.256**	1.237**	1.386**
	(.566)	(.519)	(.530)
Share of high-educated who are parents	.295	--	--
	(.533)		
Interaction (size and share high-educated)	--	--	–.120*
			(.065)
(Initial) SAMS FE:s	Yes	Yes	Yes
Ethnic group FE:s	Yes	Yes	Yes
Year of arrival FE:s	Yes	Yes	Yes
Year of graduation FE:s	Yes	Yes	Yes
Number of observations	20,039	20,039	20,039

Notes: Neighborhood characteristics are measured in logs. The sample consists of refugee immigrants whose parents arrived during the period 1987−1991 and who completed compulsory school no later than 2003. Panel A displays estimates of neighborhood characteristics measured at the year of arrival. Panel B shows the corresponding estimates for the year of graduation. All regressions control linearly for the subject's and the mother's age, and include dummies for each parent's educational attainment, family size, gender and missing values. Column (3) presents estimates where the coefficients are evaluated at the mean of the other variable. Standard errors robust for clustering at the SAMS*ethnic group level (5947 cells) in parentheses. ** = significant at 5 % level; * = significant at 10 % level

The magnitudes involved suggest that a given change in the educational attainment of the ethnic community is almost twice as important as the size of the community. However, if the estimates are evaluated at the typical variation in the data they are about as important: standard deviation changes in quality (education) and quantity (size of community) improves student performance by 0.9 percentile ranks. The effect of quantity is less precisely estimated (it is significant at the 10-percent level).[23]

Since the human capital of the parents has no additional effect on student performance, we move on to the more parsimonious specification in column (2). The size of the coefficients is reduced somewhat but the level of human capital in the ethnic community remains statistically significant at the 5-percent level.

The interaction between quantity and quality may also matter, i.e., it may be more (or less) important to have high quality peers in a sizable community. Column (3) adds the interaction of the two variables to the specification. The point estimate on the interaction is insignificant, and therefore we drop this specification from here on.

The estimates in Panel A of Table 2 are not subject to bias due to residential sorting. To illustrate the importance of sorting bias, Panel B presents results from models where the characteristics of the ethnic community are measured at the time of graduation. The results show that sorting bias is a concern for the estimate on the size of the community: the estimate is statistically significant and has the opposite sign compared to the corresponding estimate in Panel A. Sorting bias does not appear to affect the estimate on the educational composition of the ethnic community.

[23] An alternative evaluation point is the standard deviation calculated within ethnic groups across neighborhoods (see Table A2). This evaluation point produces somewhat smaller effects but does not change the relative importance of quantity and quality.

We noted in the previous section that the studied refugees became more concentrated with time in Sweden. The size of the ethnic community in the neighborhood doubles between the time of arrival and the time of graduation. The results in Table 2 imply that it is primarily less-skilled families (in the unobserved sense) that relocate to neighborhoods where ethnic concentration is higher. This pattern is similar to the findings of Edin et al. (2003), who also concluded that sorting inflicts a negative bias on the estimate on the number of peer contacts. Note that we arrive at this conclusion despite having very flexible controls for neighborhoods and regions of origin.

Of course we have subjected the baseline specification to a number of specification checks. We find no evidence suggesting that the neighborhood effects are non-linear and no evidence of substantial attenuation bias resulting from small source countries being aggregated for confidentiality reasons (see Table A3). We have also experimented with alternative outcome variables. A particularly interesting question is whether segregation influences host country skills. To shed light on this question we have run regressions where the outcome is grade in Swedish.[24] The results suggest that there is no impact of ethnic peers for Swedish grades: the estimate on the size of the community is –0.01 (with a standard error of 0.28) and the estimate on the share high educated is 0.52 (with a standard error of 0.45). The weaker effects for this particular outcome make sense and have several interpretations. If it is the human capital of the ethnic peers that matters, it is reasonable that we estimate smaller effects where adults have less to contribute; another contributing factor is that there may be weaker incentives to learn the host country language in ethnic neighborhoods.

[24] These estimates should be interpreted cautiously since immigrant students are allowed to choose between two different tracks: a standard track and a special track for immigrants. This introduces a potential selection problem; however, we find no evidence suggesting that the ethnic network affects the choice of track.

4.3.2 Analyses by subgroups

We have re-estimated the baseline model of column (2) in Table 2 for some subgroups; the results are presented in Table 3. With this heterogeneity analysis we want to shed light on two questions: (1) Are weak or strong groups primarily affected by the characteristics of the ethnic community? (2) At what age are individuals primarily susceptible to neighborhood influences?

To address the first question we present separate estimates by gender (cols. (1) and (2)) and parental education (cols. (3) and (4)). The estimates by gender show that boys (who perform poorly in school) are significantly influenced by the number of peers, whereas girls are not. A similar pattern is available in columns (3) and (4), where the size of the community has a positive and significant for children from "non-academic" families (who perform less well in school). The effects of the human capital of the ethnic community do not vary by gender and educational background.

Table 3 Differential effects with respect to background characteristics.

	By gender		By parental education		By age at immigration	
	Boy	Girl	Academic family	Non-Academic family	0–6	7–
	(1)	(2)	(3)	(4)	(5)	(6)
Size of ethnic community	1.279**	–.441	–.121	.946**	1.306**	–.087
	(.396)	(.450)	(.473)	(.454)	(.567)	(.670)
Share high educated	1.358**	1.091	1.521*	1.169*	2.222**	–.321
	(.619)	(.697)	(.892)	(.690)	(.944)	(.415)
(Initial) SAMS FE:s	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic group FE:s	Yes	Yes	Yes	Yes	Yes	Yes
Year of arrival FE:s	Yes	Yes	Yes	Yes	Yes	Yes
Year of graduation FE:s	Yes	Yes	Yes	Yes	Yes	Yes
Mean (sd) of the dependent variable	36.60 (26.86)	44.78 (28.54)	48.13 (28.52)	33.67 (25.63)	44.12 (28.37)	38.05 (27.43)
Number of observations	10,598	9,441	9,407	10,632	7,940	12,099

Notes: Neighborhood characteristics are measured in logs. The sample consists of refugee immigrants whose parents arrived during 1987−1991 and who completed compulsory school no later than 2003. Where appropriate, the regressions control linearly for the subject's and the mother's age, with dummies for each parent's educational attainment (five levels), family size, gender and missing values. Standard errors robust for clustering at the SAMS*ethnic group level (5,947 cells) in parentheses. "Academic family" is defined as having at least one parent who has completed at least university preparatory upper-secondary school. ** = significant at 5 % level; * = significant at 10 % level.

The differential effects of the size of the peer group are interesting and shed some light on the sorting pattern in our data. Boys and children with a less-educated family background perform worse than average in school. The observed determinants of school outcomes are, arguably, positively associated with the unobserved factors determining school performance. The results presented in columns (1) to (4) thus suggest that it may be beneficial for students from weak backgrounds to sort themselves into ethnic communities, which is also the sorting pattern implied by the results in Table 2.

To address the second question we split the sample by age at migration (0–6 and 7–). The estimates in columns (5) and (6) suggest that initial neighborhood characteristics are only important for children arriving before the start of compulsory school. This result has two possible interpretations. One is that skills are primarily shaped in early childhood (cf. Cunha and Heckman, 2007). A second interpretation is that the estimates to some extent reflect a cumulative effect of peer contacts: younger migrants have on average been exposed to the environment longer and thereby received a higher treatment dose. If the second interpretation applies, we would expect the estimates for the older age group to be smaller but of the same sign as for the younger age group. Since this is not the case, we favor the first interpretation: the neighborhood characteristics have a bigger impact for children who arrived at a young age because their skills are more malleable.

4.4 Longer run outcomes

Do the effects on GPA feed on to longer run educational attainment? Intuitively the answer should be “yes” since the compulsory school GPA determines admission into upper-secondary school, and the completion of upper-secondary school is a prerequisite

for entering university. Nevertheless, it may well be that the effects are too small to alter future educational outcomes. It may also be the case that the effects are concentrated to parts of the GPA distribution where variations in GPA do not substantially alter future outcomes.

The children in our data are too young to render an analysis of university entrance meaningful.[25] To address the above question we therefore examine an intermediate outcome: whether the initial neighborhood affects the probability of graduating upper-secondary school on time (i.e., by age 19 which is the normal graduation age). Graduating on time is a powerful predictor of university graduation, as argued above. A second virtue is that we can observe this outcome for all individuals in our data, since we are able to use upper-secondary graduation data through 2009.

Table 4 The effect of neighborhood characteristics on the probability of graduating from upper-secondary school by age 19

	Total sample	By gender		By parental education		By age at immigration	
		Boy	Girl	Academic family	Non-Academic family	0–6	7–
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Size of ethnic community	.007	.019**	–.008	.002	.014*	.008	.006
	(.005)	(.007)	(.008)	(.009)	(.007)	(.010)	(.007)
Share high educated	.011	.020*	.004	–.002	.023*	.033*	.002
	(.008)	(.012)	(.013)	(.016)	(.012)	(.017)	(.012)
(Initial) SAMS FE:s	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic group FE:s	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year of arrival FE:s	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year of graduation FE:s	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean (sd) of the	.426	.380	.477	.516	.346	.530	.357
dependent variable	(.494)	(.485)	(.500)	(.500)	(.476)	(.500)	(.479)
Number of observations	20,039	10,598	9,441	9,407	10,632	7,940	12,099

Notes: Neighborhood characteristics are measured in logs. The sample consists of refugee immigrants whose parents arrived during 1987−1991 and who completed compulsory school no later than 2003. Graduation from upper-secondary school is observed through 2009. Where appropriate, the regressions control linearly for the subject's and the mother's age, and include dummies for each parent's educational attainment (five levels), family size, gender and missing values. Standard errors robust for clustering at the SAMS*ethnic group level in parentheses. "Academic family" is defined as having at least one parent who has completed at least university preparatory upper-secondary school. ** = significant at 5 % level; * = significant at 10 % level.

[25] The median age at entry in Sweden was 22.8 among university entrants in 2004 (OECD 2006b).

Table 4 reports estimates for the entire population as well as for the sub-groups that we analyzed in Table 3. On average, the effects are weaker for this longer run outcome than for compulsory school GPA. But the results suggest that there are effects for the disadvantaged subgroups whose GPA was affected: we find effects for boys but not for girls, and for children with less educated parents but not for children with high-educated parents. For children who migrated at a young age we find that the share of high-educated (but not the size of the community) has an impact on educational attainment.

How big are these effects? Take boys as an example. A standard deviation increase in the size of the community (share high educated) raises the probability of graduating from upper-secondary school on time by 2.7 (1.5) percentage points. This magnitude corresponds to a relative increase of 7.2 (4.0) percent or 10.0 (5.6) percent of the performance gap between immigrant and native boys born in 1981.[26]

4.5 IV estimates

The estimates we have presented so far are estimates of well-defined causal effects: the effects of initial exposure to a neighborhood with certain characteristics. But they do not explicitly speak to the question of why the initial neighborhood is of importance. The initial neighborhood may have an independent effect on its own, but also because it predicts future neighborhood characteristics.

To investigate the latter possibility we estimate a model assuming that average exposure to neighborhood characteristics has an effect on educational outcomes, using initial characteristics as instruments.[27] The critical assumption is then that initial

[26] For the 1981 birth cohort, the probability of graduating from upper-secondary school on time was 63.9% for native-born boys and 37.1% for immigrant boys.
[27] A more reasonable hypothesis is that the entire sequence of neighborhood characteristics has an impact on school achievement. However, we do not have a sufficient number of instruments to identify such a model.

exposure is excludable from the outcome equations, conditional on average exposure between the time of arrival and the time of outcome measurement. This holds if there is no impact of the initial location or if its impact equals that of average exposure.

We think there is reason to expect that initial exposure to neighborhood characteristics is more important than later exposure (which means that initial exposure is non-excludable). This is an implication of a model where "skill-begets-skill" (e.g., Cunha and Heckman 2007); then initial conditions shape the future accumulation of human capital. The differential effects by arrival age documented in Tables 3 and 4 are in line with this hypothesis. If initial exposure is more important than average exposure we expect an upward bias of the IV estimates.

In our view, the main advantage of the IV estimates is that they implicitly correct for mobility between the time of arrival and the time of measurement. This is especially useful when it comes to estimating differential effects across groups. For instance, the differential effects by parental education may be due to higher mobility out of the initial neighborhood among well-educated families.

Table 5 presents the IV (2SLS) estimates. We restrict attention to the percentile ranked GPA, since we do not have the data to measure neighborhood characteristics after 2003. In the interest of conserving space, we do not report the first-stage relationships. These estimates (available upon request) suggest that the instruments have substantial predictive power in all columns; weak instruments do not plague our estimates.

With the above-mentioned caveats in mind, we note that the IV estimates exhibit the same pattern as above. They are much larger, however, which is due to the fact that

some 75 percent of the families move out of the initial neighborhood. Evaluated at the same point as the estimates in Table 2, the estimate in column (1) implies that a greater share of high-educated peers improves performance by 3.5 percentile ranks. Although potentially representing an upper bound of the effect of average exposure to a highly educated ethnic community, the magnitude is not implausibly large: it corresponds to a quarter of the performance difference between children with mothers who have a university degree as opposed to a compulsory school degree. The estimates also support our earlier conclusions on heterogeneous neighborhood effects. The correction for differential mobility rates, however, increases the relative importance of the share of high-educated in the ethnic community for children with an academic background.

Table 5 IV estimates of the effect of average exposure to neighborhood characteristics on compulsory school GPA

	Total sample	By gender		By parental education		By age at immigration	
		Boy	Girl	Academic family	Non-Academic family	0–6	7–
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Size of ethnic community	1.097	3.069**	–.586	–.474	1.872*	3.310**	–.279
	(.701)	(1.054)	(1.043)	(1.189)	(.985)	(1.624)	(.800)
Share high educated	4.683*	5.458*	3.291	7.714*	3.207	11.941**	.285
	(2.462)	(2.916)	(3.745)	(4.641)	(2.694)	(4.681)	(2.604)
(Initial) SAMS FE:s	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic group FE:s	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year of arrival FE:s	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year of graduation FE:s	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean (sd) of the dependent variable	40.45	36.60	44.78	48.13	33.67	44.12	38.05
	(27.96)	(26.86)	(28.54)	(28.52)	(25.63)	(28.37)	(27.43)
Number of observations	20,039	10,598	9,441	9,407	10,632	7,940	12,099

Notes: 2SLS estimates. Neighborhood characteristics are measured in logs and averaged over the observation period from year of arrival to graduation from compulsory school. The sample consists of refugee immigrants whose parents arrived during 1987−1991 and who completed compulsory school no later than 2003. The regressions control linearly for the subject's and the mother's age, and include dummies for each parent's educational attainment, family size, gender and missing values. Standard errors robust for clustering at the SAMS*ethnic group level in parentheses. ** = significant at 5 % level; * = significant at 10 % level.

5 Concluding remarks

This paper examines whether the size and characteristics of the ethnic community affect school performance of immigrant children in Sweden. To handle sorting in the

residential market, the analysis uses a governmental refugee placement policy in place in the late 1980s and early 1990s.

The results show that peers matter. The number of highly educated in the local ethnic community has a positive effect on compulsory school grades. Separating this effect into its components, we find that a higher level of education among fellow countrymen in the assigned neighborhood has a positive effect: A standard deviation increase in the fraction of highly educated peers raises student performance by 0.9 percentile ranks. A standard deviation increase in the size of the ethnic community has about the same effect, but the effect is less precisely estimated.

We have also presented some evidence on the importance of handling the problems associated with residential sorting in studies relating contextual variables to individual outcomes. Like some previous studies on adult migrants (Edin et al. 2003, Åslund and Fredriksson 2009), we find that one is likely to infer—erroneously—that the number of peer contacts has a negative effect on school performance if sorting bias is not addressed appropriately. In this respect, our analysis of heterogeneous effects reveals an interesting pattern. Disadvantaged students gain more by having many peers around than other students. And it is also families with disadvantaged students that move to ethnically concentrated areas. The sorting pattern thus appears to be rational from the point of view of the disadvantaged groups.

Our evidence also suggests that the effect on the GPA will translate into improvements in educational attainment for some groups. For boys and children in less-educated families, we find that the probability of graduating from upper-secondary school on time—a strong predictor of later obtaining a university degree—is increasing

in the size of the ethnic community as well as in educational attainment in the ethnic group. These two groups share two features: average performance is poor (in compulsory as well as upper secondary education), and the ethnic community has a relatively strong impact on compulsory school performance.

Another general finding is that children who migrated at a young age are more susceptible to peer influences than older child migrants. The characteristics of the neighborhood community thus appear to be more important for children who are in their formative years.

Overall we view the estimates of heterogeneous effects as being remarkably consistent with the effects of educational interventions on pupil performance. One example is the literature showing that reductions in class size tend to have more positive effects for disadvantaged and younger children (e.g., Krueger 1999; Robinson 1990).

Are the neighborhood effects small or large? They may seem small relative to the importance of individual or family characteristics. For instance, the effect of a standard deviation increase in the share high-educated in the assigned neighborhood corresponds to 10 percent of the grade difference between refugee immigrants and the native-born, and 6 percent of the attainment difference between native-born boys and immigrant boys. But relating the estimates to individual or family characteristics is hardly the right comparison. A better comparison is to educational interventions, such as the above-mentioned case of variations in class size. A comparison of our estimates to those from the class size experiment in STAR (see Krueger 1999), suggest that the effects we

estimate are about half as large as the effect of variations in class size.[28] It seems to us that this is a rather large effect.

How do our results relate to the previous literature on immigrants and ethnic minorities? Let us first note that the size and characteristics of the ethnic enclaves have been found to be important for refugee immigrants in other contexts than the Swedish one: Denmark (Damm 2009b) and the U.S. (Beaman 2009) are two examples. Second, we think our results are quite consistent with studies credibly identifying the importance of the neighborhood for immigrant outcomes. On our reading, the major result in this literature is that the quality of the neighborhood (or school) has a positive impact on immigrant outcomes. For instance, Gould et al. (2004) find that Ethiopian immigrants to Israel who were assigned to a high-quality school did better in school; Bertrand et al. (2000) find that being exposed to a greater number welfare recipients (a reduction in quality) increased the individual probability of being on welfare (a negative outcome).[29] Our main result is that the share of highly educated in the community has a positive impact on student performance for practically all groups we have examined (child migrants who arrive after compulsory school start is the only exception). An increase in the size of the community improves performance primarily for groups expected (and

[28] To arrive at this conclusion we did the following calculation. According to Krueger's estimates, the effect of being randomized to a small class in Kindergarten on student achievement in grade 3 is 5.6 percentile ranks for the average student. Dividing this estimate by the difference in class size in small and regular sized classes (7.3 pupils) and multiplying by the standard deviation of class size in regular sized classes (2.21), we conclude that a standard deviation reduction of class size improves student performance by 1.7 percentile ranks. An analogous calculation for black students (presumably a more relevant comparison group) gives an effect size of 2.4 percentile ranks. These two effect sizes should be compared to the effect size of 0.9 that we estimate here.

[29] The papers by Edin et al (2003), Beaman (2009), Åslund and Fredriksson (2009) also belong to this category. For instance, Beaman (2009) finds that, on average, the size of the ethnic community has no impact on the employment probability. But an increase in the number of fellow countrymen who has been in the U.S. for at least two years has a positive impact on the employment probability. Beaman attributes this result to the fact that individuals who have been in the host country for some time provides information on job contacts (which represents an improvement in the quality of the network in our terminology).

observed) to do poorly in school. If peer quality relative to individual performance is relevant, this is in line with the literature emphasizing the quality of the neighborhood.

At first glance, our main result is not in line with studies examining how minority (black) students are affected by desegregation policies. The typical result is that desegregation improves school outcomes for blacks. However, it is not clear that the results for African-Americans translate to immigrants; they are obviously different groups and to some extent face different problems. Moreover, desegregation implies two things: less exposure to the own group and a change in peer characteristics. Therefore, estimates of the effects of desegregation policies answer a different question than the one we attempt to answer in this paper. Note, finally, that we have presented tentative results suggesting that, in contrast to the positive influence from ethnic peers, an immigrant-dense environment has a negative impact on student performance.

What mechanisms underlie our results? This question is very interesting, but the kind of register data we are using are not well-suited for answering it. A very fruitful exercise would be to merge register data with survey data to try to pin down the mechanisms (see Lavy et al. 2008 for a recent example). This opportunity is unavailable to us. The pattern of our results offers some insights, however. For instance, we find that the characteristics of the entire ethnic community are as important as the characteristics of the children in that community. This may suggest that, in a tightly defined neighborhood, all adults serve as role models. This interpretation is of course somewhat speculative. A thorough understanding of the mechanisms is a highly relevant topic for future studies.

References

Angrist, Joshua D., and Kevin Lang. 2004. "Does School Integration Generate Peer Effects? Evidence from Boston's Metco program.", *American Economic Review*, 94(5): 1613–1634.

Andersson Roger. 2000. "Etnisk och socioekonomisk segregation i Sverige 1990–1998." In *Välfärdens förutsättningar,* SOU 2000:37, ed. Johan Fritzell. Stockholm: Fritzes.

Åslund, Olof, Per-Anders Edin, Peter Fredriksson, and Hans Grönqvist (2009), "Peers, Neighborhoods, and Immigrant Student Achievement – Evidence from a Placement Policy." Working Paper 2009:20, IFAU

Åslund, Olof, and Peter Fredriksson. 2009. "Peer Effects in Welfare Dependence—Quasi-experimental Evidence." *Journal of Human* Resources, 44(3): 799–825.

Åslund, Olof, and Hans Grönqvist. 2010. "Family Size and Child Outcomes: Is there Really no Trade-off?" *Labour Economics*, 17(1): 130–139.

Åslund, Olof, John Östh, and Yves Zenou. 2010. "How Important is Access to Jobs? Old Question—Improved Answer." *Journal of Economic Geography*, 10(3): 389–422.

Åslund, Olof, and Dan-Olof Rooth. 2007. "Do When and Where Matter? Initial Labor Market Conditions and Immigrant Earnings." *Economic Journal*, 117(518): 422–448.

Beaman, Lori A. (2009). "Social Networks and the Dynamics of Labor Market Outcomes: Evidence from Refugees Resettled in the U.S." Manuscript, Northwestern University.

Bertrand, Marianne, Erzo F. P. Luttmer, and Sendhil Mullainathan. 2000. "Network Effects and Welfare Cultures." *Quarterly Journal of Economics*, 115(3): 1019–1055.

Björklund, Anders, Melissa, A. Clark, Per-Anders Edin, Peter Fredriksson, and Alan B. Krueger. 2005. *The Market Comes to Education in Sweden—An Evaluation of Sweden's Surprising School Reforms*. New York: Russell Sage Foundation.

Bleakley, Hoyt and Aimee Chen. 2004. "Language Skills and Earnings: Evidence from Childhood Migrants." *Review of Economics and Statistics*, 86(2): 481–496.

Böhlmark, Anders. 2008. "Age at Immigration and School Performance: A Siblings Analysis using Swedish Data." *Labour Economics*, 15(6): 1366–1387.

Boozer, Michael A., Alan B. Krueger, and Shari Wolkon. 1992. "Race and School Quality since Brown v. Board of Education." *Brookings Papers on Economic Activity*, Microeconomics: 269–326.

Borjas, George J. 1995. "Ethnicity, Neighborhoods, and Human-capital Externalities." *American Economic Review*, 85(3): 365–390.

Bygren, Magnus, and Ryszard Szulkin. 2007. "Ethnic Environment during Childhood and the Educational Attainment of Immigrant Children in Sweden." Working Paper 2007:8, SULCIS, Stockholm University.

Card, David, and Jesse Rothstein. 2007. "Racial Segregation and the Black-white Test Score Gap." *Journal of Public Economics*, 91(11–12): 2158–2184.

Cunha, Flavio, and James Heckman. 2007. "The Technology of Skill Formation." *American Economic Review*, 97(2), 31–47.

Cliffordson, Christina. 2004. "Betygsinflationen i de målrelaterade gymnasiebetygen." *Pedagogisk forskning*, 9: 1–14.

Corak, Miles. 2006. "Do Poor Children Become Poor Adults? Lessons from a Cross Country Comparison of Generational Earnings Mobility." IZA Discussion Paper 1993.

Cortes, Kalena E. 2006. "The Effects of Age at Arrival and Enclave Schools on the Academic Performance of Immigrant Children." *Economics of Education Review*, 25(2): 121–132.

Cutler, David M., and Edward L. Glaeser. 1997. "Are Ghettos Good or Bad?" *Quarterly Journal of Economics*, 112(3), 827–872.

Damm, Anna P. 2009a. "Determinants of Recent Immigrants' Location Choices: Quasi-Experimental Evidence." *Journal of Population Economics*, 22(1): 145–174.

Damm, Anna P. 2009b. "Ethnic Enclaves and Immigrant Labor Market Outcomes: Quasi-Experimental Evidence." *Journal of Labor Economics*, 27(2): 281–314.

Edin, Per-Anders, Peter Fredriksson, and Olof Åslund. 2003. "Ethnic Enclaves and the Economic Success of Immigrants: Evidence from a Natural Experiment." *Quarterly Journal of Economics*, 118(1): 329–357.

Edin, Per-Anders, Peter Fredriksson, and Olof Åslund. 2004. "Settlement Policies and the Economic Success of Immigrants." *Journal of Population Economics*, 17(1): 133–155.

Goel, Deepti., and Kevin Lang. 2009. "The Role of Social Ties in the Job Search of Recent Immigrants." CLSRN Working Papers 2009-12, UBC Department of Economics.

Gould, Eric D., Victor. Lavy, and M. Daniele Paserman. 2004. "Immigrating to Opportunity: Estimating the Effect of School Quality using a Natural Experiment on Ethiopians in Israel." *Quarterly Journal of Economics*, 119(2): 489–526.

Grönqvist, Hans. 2006. "Ethnic Enclaves and the Attainments of Immigrant Children." *European Sociological Review*, 22(4): 369–382.

Grogger, Jeff. 1996. "Does School Quality Explain the Recent Black/white Wage Trend?" *Journal of Labor Economics*, 14(2), 231–253.

Guryan, Jonathan. 2004. "Desegregation and Black Dropout Rates." *American Economic Review*, 94(4), 919–943.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2009. "New Evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement." *Journal of Labor Economics*, .27(3), 349–383.

Hoxby, Caroline. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." NBER Working Paper 7867.

Jensen, Peter, and Astrid W. Rasmussen. 2008. "Immigrant and Native children's Cognitive Outcomes and the Effect of Ethnic Concentration in Danish Schools." Manuscript, University of Aarhus.

Kling, Jeffrey, Jeffrey Liebman, and Lawrence Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75(1): 83-119.

Krueger, Alan B. 1999. "Experimental Evidence of Education Production Functions." *Quarterly Journal of Economics*, 114(2): 497–532.

Lavy, Victor, M. Daniele Paserman, and Analia Schlosser. 2008. "Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom." NBER Working Paper 14415.

OECD. 2006a. *Where Immigrants Succeed: A Comparative Review of Performance and Engagement in PISA 2003*. Paris: Organization for Economic Cooperation and Development.

OECD. 2006b. *Education at a Glance*. Paris: Organization for Economic Cooperation and Development.

OECD. 2007. *Jobs for Immigrants: Labour Market Integration in Australia, Denmark, Germany and Sweden*, Organization for Economic Cooperation and Development.

Oreopoulos, Philip. 2003. "The Long-run Consequences of Living in a Poor Neighborhood." *Quarterly Journal of Economics*, 118(4): 1533–1575.

Rivkin, Steven G. 2000. "School Desegregation, Academic Attainment, and Earnings." *Journal of Human Resources*, 35(2): 333–346.

Robinson, Glen E. 1990. "Synthesis of Research on the Effects of Class Size." *Educational Leadership*, 47(7): 80–90.

Stark, Oded. 1991. *The Migration of Labor*, Oxford: Blackwell.

Appendix

Table A1 The correlation between neighborhood characteristics and with pre-determined parental and child characteristics

	(log) Size ethnic community		(log) Share high educated	
	(1) Year of arrival	(2) Year of graduation	(3) Year of arrival	(4) Year of graduation
***Mother characteristics*:**				
Age	.003*	.008**	–.011	.008
	(.002)	(.003)	(.009)	(.007)
Education:				
High school ≤ 2 years	–.047	–.143**	–.092	–.190
	(.030)	(.051)	(.151)	(.128)
High school > 2 years	–.033	–.179**	–.239	–.311**
	(.030)	(.052)	(.152)	(.128)
University ≤ 2 years	–.029	–.254**	.119	–.513**
	(.034)	(.059)	(.170)	(.147)
University > 2 years	–.028	–.347**	–.061	–.819**
	(.040)	(.068)	(.203)	(.187)
***Father characteristics*:**				
Education:				
High school ≤ 2 years	.001	–.211**	.097	–.099
	(.033)	(.056)	(.175)	(.144)
High school > 2 years	–.009	–.100**	–.091	–.148
	(.033)	(.057)	(.170)	(.133)
University ≤ 2 years	–.053	–.188**	–.243	–.241
	(.038)	(.066)	(.187)	(.154)
University > 2 years	–.074**	–.280**	–.227	–.179
	(.037)	(.065)	(.190)	(.153)
Child characteristics:				
Female	–.026**	.037*	.043	.102*
	(.429)	(.028)	(.059)	(.060)
Age at immigration	.004	.009	.040	–.000
	(.015)	(.028)	(.083)	(.077)
Family size FE:s	Yes	Yes	Yes	Yes
(Initial) SAMS FE:s	Yes	Yes	Yes	Yes
Ethnic group FE:s	Yes	Yes	Yes	Yes
Year of arrival FE:s	Yes	Yes	Yes	Yes
Year of grad. FE:s	Yes	Yes	Yes	Yes

Notes: Estimates on individual characteristics for the specification in Table 1, column (1). Number of observations is 20,039. The sample consists of refugee immigrants whose parents arrived during the period 1987−1991 and completed compulsory school not later than 2003. Standard errors are robust for clustering at the SAMS*ethnic group level (5947 cells) in parentheses. ** = significant at 5 % level; * = significant at 10 % level.

Table A2 Summary statistics

Variable	Mean	Standard deviation
Subject:		
GPA (percentile rank)	40.45	27.96
Graduating from upper-secondary school at age≤19	.43	.49
Age (in 2003)	21.95	3.84
Age at immigration	8.00	3.8
Female	.47	.50
Sibship size	2.99	1.56
***Mother*:**		
Age (in 2003)	47.38	6.39
Education: Compulsory school	.50	.50
Upper secondary school ≤ 2 years	.14	.34
Upper secondary school > 2 years	.17	.38
University ≤ 2 years	.11	.31
University > 2 years	.08	.28
***Father*:**		
Age (in 2003)	51.48	6.99
Education: Compulsory school	.42	.49
Upper secondary school ≤ 2 years	.14	.35
Upper secondary school > 2 years,	.17	.38
University ≤ 2 years	.12	.33
University > 2 years	.15	.35
Regional characteristics: Year of arrival		
Share high-educated in own group	34%	
Share high-educated in immigrant group	31%	
"Ethnic" concentration	1.6%	
Immigrant concentration	19%	
Population size	1528	
ln(share high-educated in own group)	–1.016	.758*
		[0.520]
ln(size of ethnic community)	2.372	1.445*
		[1.100]
ln(size of immigrant community)	4.830	1.217*
		[0.769]
Regional characteristics: Year of graduation		
Share high-educated in own group	39%	
Share high-educated in immigrant group	38%	
"Ethnic" concentration	3.2%	
Immigrant concentration	31%	
Population size	2012	

Notes: The regional characteristics are defined with respect to the adult population aged 25-65. Summary statistics for each parent's educational attainment is conditional on having found this information in the records. * The standard deviations are calculated excluding "empty cells", i.e., excluding the observations where there is no other immigrant from the same source country in the neighborhood. The standard deviations within square brackets correspond to the standard deviation within ethnic group across neighborhoods.

Table A3 Region of birth

Region of birth	Percent of sample
1. Former Yugoslavia	5.2
2. Poland	5.5
3. The Baltic states (Estonia, Latvia, Lithuania)	0.3
4. Eastern Europe 1 (Rumania, The former USSR, Bulgaria, Albania)	6.0
5. Eastern Europe 2 (Hungary, The former Czechoslovakia)	2.4
6. Mexico and Central America (El Salvador, Mexico Other countries)	1.6
7. Chile	13.3
8. Other South America (Peru, Brazil, Colombia, Argentina, Uruguay, Other countries)	2.0
9. African Horn (Ethiopia, Somalia, Sudan, Djibouti)	5.0
10. North Africa (Arabic countries) and Middle East (Lebanon, Syria, Morocco, Tunisia, Egypt, Algeria, Israel, Palestine, Jordan, Other countries)	17.8
11. Other Africa (Gambia, Uganda, Zaire Ghana, Other countries)	1.1
12. Iran	25.5
13. Iraq	4.8
14. Turkey	3.8
15. South East Asia (Vietnam, Thailand, the Philippines, Malaysia, Laos Other countries)	3.9
16. Other Asia (Sri Lanka, Bangladesh, India, Afghanistan, Pakistan)	1.7
Total	100

Table A4 Estimates on other characteristics for specification in Table 2, column (2)

	Dependent variable: Percentile ranked GPA
Child characteristics:	
Female	8.137**
	(.371)
Age at immigration	–4.694**
	(.429)
***Mother characteristics*:**	
Age	.124**
	(.040)
Education: Compulsory school	--
High school ≤ 2 years	4.716**
	(.800)
High school > 2 years	5.886**
	(.732)
University ≤ 2 years	11.339**
	(.897)
University > 2 years	13.561**
	(1.039)
Missing education	.729
	(.939)
***Father characteristics*:**	
Missing father	1.237
	(1.057)
Education: Compulsory school	--
High school ≤ 2 years	3.475**
	(.848)
High school > 2 years	3.443**
	(.792)
University ≤ 2 years	8.061**
	(.880)
University > 2 years	11.697**
	(.905)
Missing education	–1.865**
	(.932)
Family size FE:s	Yes
(Initial) Municipality FE:s	Yes
Ethnic group FE:s	Yes
Year of arrival FE:s	Yes
Year of graduation FE:s	Yes
Number of observations	20,039
R-squared	0.335

Notes: Estimates on individual characteristics for the specification in Table 1, column (1). The sample consists of refugee immigrants whose parents arrived during the period 1987−1991 and completed compulsory school not later than 2003. The regression also controls for the regional characteristics listed in Table 1, column (1) and indicator variables controlling for the SAMS*(ethnic group) "cell" having no observations. Standard errors are robust for clustering at the SAMS*ethnic group level (5947 cells) in parentheses. ** = significant at 5 % level; * = significant at 10 % level.