

Does School Segregation Lead to Poor Educational Outcomes? Evidence from Fifteen Cohorts of Swedish Ninth Graders

Maria Brandén

Institutet för analytisk sociologi, Linköpings Universitet och Sociologiska Institutionen, Stockholms Universitet

Gunn Elisabeth Birkelund

Institutet för analytisk sociologi, Linköpings Universitet och Institutt for sociologi og samfunnsgeografi, Oslo Universitet

Ryszard Szulkin

Sociologiska Institutionen, Stockholms Universitet

Does School Segregation Lead to Poor Educational Outcomes?

Evidence from Fifteen Cohorts of Swedish Ninth Graders

ABSTRACT

We examine the impact of ethnic school segregation on the educational outcomes of students, using Swedish population register data. Through a school fixed effects, family fixed effects, and a two-way school- and family fixed effects design, we adjust for selection effects related to variation in the student composition across schools. The analyses show that students' grades are relatively unaffected by the proportion of immigrant schoolmates. However, it has a small negative effect on levels of eligibility for upper secondary school. Furthermore, immigrants' educational outcomes are weakly positively affected by the proportion of peers with the same national background as themselves.

INTRODUCTION

Over recent decades, many European countries, including Sweden, have experienced a rapid inflow of immigrants, accompanied by an increase in ethnic residential and school segregation (Biterman and Franzén 2007; Musterd 2005; Böhlmark, Holmlund and Lindahl 2015).

Together, these processes have changed the demographic composition of the social spaces in which children spend most of their time during their most formative years. In Sweden, there has been a simultaneous increase in ethnic school segregation and in the between-school variation in student performance (Holmlund et al. 2014). An intuitive conclusion drawn by policy makers, school administrators, journalists and parents alike is that ethnic school segregation impedes the educational progress and future opportunities of both native born and immigrant students. Ethnic school segregation is therefore a cause of considerable concern.

The purpose of this paper is to analyze whether these concerns are justified by studying how ethnic school segregation in lower secondary schools influences the educational outcomes of students in their final year of compulsory education.

The performance of the education system is a critical factor in integrating immigrants and the children of immigrants (Alba, Sloan and Sperling 2011), and an extensive research literature has explored the association between the proportion of immigrants in schools and their students' educational outcomes (Hermansen and Birkelund, 2015; Hardoy and Schøne 2013; Geay, McNally, and Telhaj 2013; Ohinata and van Ours 2013; Brunello and Rocco 2013; Bifulco, Flechter, and Ross, 2011; Jensen and Rasmussen 2011; Hanushek, Kain and Rivkin 2009; Gould, Lavy, and Paserman 2009; Cebolla-Boado 2007; Fekjær and Birkelund 2007; Szulkin and Jonsson 2007). One overarching finding from these studies is that there is a negative correlation between the average performance of students and the proportion of immigrants or children of immigrants. However, when researchers have attempted to identify the causal effects of having a high proportion of immigrants or children of immigrants in the

school environment, findings have been mixed. This is probably due to differences in methodology, country and period of data collection.

One major analytical challenge for all studies that attempt to identify macro-level effects on micro-level outcomes is the selection problem (Manski 1993). Children are not randomly allocated to schools and non-native students are more likely to attend schools with disadvantaged native students. Variation in student outcomes between different schools may be the result of three different types of processes: environmental effects, selection effects, and social interaction effects. The differences between these processes can be described as follows: “An environmental effect is operative if we do what we do because we are where we are. A selection effect is operative if we do what we do because we are who we are. And finally, a social-interaction effect exists if we do what we do because others do what they do” (Åberg and Hedström 2011:205).

The aim of the present study is to empirically separate these different processes and to estimate the strength of the latter effect in schools with varying proportions of immigrant children. Our analytical strategy, described in detail below, is to net out the impact originating from school resources and the social selection of students to schools by means of the step-wise introduction of school fixed effects, family fixed effects, and the use of a two-way school- and family fixed effects design. In this way we contribute to previous research by exploiting variation over time within schools and within families (sibling pairs). This design adjusts for all time-invariant school and family characteristics, which should substantially diminish the selection problems mentioned above. Separating out individual effects (the selection of students with particular individual characteristics) and environmental effects (school characteristics, such as reputation, teacher quality and other resources) creates an exceptional opportunity for identifying causal social interaction effects on the relationship between the concentration of immigrants in schools and the students’ educational outcomes at

the end of lower secondary school. We analyze an unusually rich Swedish data set which includes all 1.5 million students who left lower secondary schools during the period 1998-2012, addressing two important educational outcomes at the conclusion of lower secondary school: students' grade point scores, and whether or not they are eligible for upper secondary school; and we examine how these outcomes are dependent on the immigrant composition of the school.

THEORY AND RECENT RESEARCH

The family of origin is the primary arena in which the socialization of children occurs. Parental educational, cultural and economic resources are important factors underlying the social inheritance of privilege in the educational system. This is why a large part of the educational disadvantage of the children of immigrants can be attributed to family background and parental characteristics such as formal education and labor market position (e.g., Jackson, Jonsson and Rudolphi 2012; Jonsson and Rudolphi 2011; Birkelund and Mastekaasa 2009; Heath, Rethon and Kilpi 2008; Kao and Thompson 2003).

However, other social spaces in which children spend their time and meet adults and peers of the same age may also be considered to constitute important arenas for socialization. Thus over and above the characteristics of children's family backgrounds, the characteristics of schools and neighborhoods may also have an impact on future life chances. By choosing schools, parents decide which peers their children meet at school. *Peer groups* are important for identity formation, and they may also influence students' future dispositions towards learning (Lavecchia et al. 2014). Children interact with each other both inside and outside the classroom, and the aspirations and behaviors of students at school may be influenced by the aspirations and behavior of their peers. Most students care about the behaviors of significant others and do not want to deviate from the "local" social norms developed in a peer group

(Akerlof and Kranton 2002). Thus the time and effort peers spend on schoolwork, how they value learning, their school attendance and their educational aspirations all constitute important parts of the learning environment (e.g. Hermansen and Birkelund 2015; Szulkin and Jonsson 2007; Winston and Zimmerman 2004).

Sociological theories of role models, collective socialization, and peer processes suggest that the concentration of disadvantaged families in neighborhoods and schools will have negative consequences for children's school achievement. Although there are some common mechanisms that may explain why segregation depresses educational achievement, one crucial question is why we should also expect effects stemming from ethnic concentration *per se*. The focus of the current paper is therefore directed at uncovering the effect of schools' ethnic composition on the school progress of students.

The spatial segregation of immigrants is in part a 'population shock' phenomenon, due to (sometimes large numbers of) newcomers tending to cluster in separate, often underprivileged, residential areas (e.g., Jensen 2001). There are good reasons to believe that such ethnic concentration produces communities in which disadvantage may be maintained across generations (e.g., Borjas 1995; Portes and Rumbaut 2001). There are several reasons for such an expectation.

First, a lack of proficiency in the language of instruction may disturb the learning environment for students. If a relatively large fraction of the students need extra attention and additional instruction, this may disrupt the classroom learning environment for all of those involved, that is, not only for the immigrant students.

Second, the probability of exposure to various forms of trauma is high among the more recent cohorts of immigrants to Sweden, since many of them are refugees from countries with a recent history of conflicts, war and persecution. The presence of immigrant students with traumatic experiences may produce social and psychological problems at school and may

influence other students who were not themselves exposed to these traumas (cf. Hällsten, Sarnecki and Szulkin 2013; Szulkin and Jonsson 2007).

Third, teachers may also raise – or lower – their expectations of the performance of the students in class depending on what they perceive to be the average ‘capacity’ of the students. The presence of traumatized or low-achieving students with special needs, such as extra language tuition, will require additional attention on the part of teachers, which may reduce the time teachers have for attending to other students (Fletcher 2010).

Forth, other adults, such as schoolmates’ parents, may also become important *role models*, transferring skills and exercising social control (e.g., Crowder and South 2003; Sampson, Morenoff, and Gannon-Rowley 2002; Coleman and Hoffer 1987). Students in schools with a high concentration of children of immigrant background are provided with many examples of adults whose educational merits from their countries of origin do not lead to adequate employment in the new country. Low levels of returns to education in the parental generation may create an atmosphere in which students underestimate the value of education and formal merits and lower their educational aspirations (cf. Morgan 2005; Zhou 1997). A more drastic version of this hypothesis states that minority groups who have long lived in marginalized circumstances may be more disposed to develop “oppositional cultures” which question the central social values of the majority society (Fordham and Ogbu 1986).

All these processes highlight the sociological relevance of social interaction effects and peer influence. However, institutional effects may also be of great relevance. The concentration of immigrant students in a school may be negatively correlated with *school quality* and school reputation. Schools with a large proportion of immigrant students, schools with many students from low-income families, or schools located in areas with a high occurrence of social problems might experience difficulties in recruiting and retaining skilled teachers (e.g., Clotfelter, Ladd and Vigdor 2005; Hanushek, Kain and Rivkin 2004). However, if schools

with high proportions of minority students or students from low-income families receive extra support from the school authorities (as in Sweden) this may counteract the negative institutional effects of segregation.

All these arguments point in the same direction: A high concentration of immigrant students in schools may have a negative impact on children's educational outcomes. This applies particularly to schools with a high concentration of relatively newly arrived immigrants. It is reasonable to assume that the expected (negative) effects will be particularly noticeable for children born outside Sweden, yet one might also expect spill-over effects for children of native Swedes and second generation immigrants.

Positive or negative peer effects on educational outcomes?

As described above, empirical studies have established that, at the school level, the *association* between high concentrations of immigrant children and low average grades is undisputable. It therefore came as a surprise – not least to researchers – that this association more or less disappeared in the context of more advanced multivariate analyses. Strong associations between school-related ethnic segregation and children's educational outcomes dissolved into weak negative regression coefficients, or no significant effects. In particular, most attempts to measure *causal* peer effects of school segregation, for instance by means of instrumental variable analysis or within-school-analysis, have identified only weak effects (Erikson 1994; Robertson and Symons 2003; Hanushek, Kain and Rivkin 2002, 2004; Szulkin and Jonsson 2007; Fekjær and Birkelund 2007; Åslund et al. 2011; Nordin 2013; Hermansen and Birkelund 2015; Ohinata and Van Ours 2013; Geay, McNally and Telhaj 2013; Jensen and Rasmussen 2011; Schwartz and Stiefel 2011; Cebolla-Boado 2007). In Israel, Gould, Lavy and Paserman (2009) showed that a 10 percent increase in the proportion of immigrants lowered the probability of passing high school matriculation exams by 1.8 percentage points. In Norway, Hardoy and Schøne (2013) showed that a 10 percent increase in the proportion of

immigrants lowered the probability of completing upper secondary school by 2 percentage points.

Clearly, immigrant peer effects *may* be negative, but there must be more to the story. Some researchers have adopted more positive interpretations, emphasizing mechanisms such as “immigrant drive” and the development of social and ethnic capital among students at school. The first explanation relates to the selection of students to schools. Migrants are usually positively selected; they are often resourceful individuals with a *drive* for success, striving for a better life in their new country of residence (Lee and Zhou 2015; Birkelund and Mastekaasa 2009; Parsons and Smeeding 2006). Most immigrants to Sweden from non-Western countries, including immigrants arriving for family reunification, may be lacking in human capital and economic resources, yet they may still differ in non-trivial, unobserved ways from low-status native parents. According to the family mobilization hypothesis, migrants transmit their drive, both genetically and socially, to their children (Parsons and Smeeding 2006), and this immigrant *drive* would lead us to expect weaker compositional effects of the socioeconomic background of immigrant students on the educational outcomes of all students. Thus, schoolmates can contribute to a productive learning environment in the classroom (Lazear 2001). These positive effects of migration may also, at least partly, counteract the negative effects of immigrants’ low socio-economic resources.

The ethnic capital argument relates to what we have labelled social interaction effects. Positive peer effects may arise at schools with a high concentration of immigrant students. In particular, close social relationships are generally more likely to develop in schools with a high proportion of individuals of *similar* ethnic or national background (Lee and Zhou 2015; Nekby 2011; McPherson, Smith-Lovin and Cook 2001; Moody 2001; Akerlof 1997; Allport 1954), and a shared ethnic identity may constitute a significant resource network (Portes, Fernandez-Kelly and Haller 2005). According to attachment theory (Coleman 1988;

Agirdag et al. 2012), supportive and close social relationships can generate social capital that positively contributes to students' well-being, school involvement, and educational outcomes. Consequently, a high concentration of immigrant students of the same ethnic background could create *positive* conditions for learning – at least for the immigrant students themselves (e.g. Fleischmann et al. 2012; Åslund et al. 2011, Bygren and Szulkin 2010).

The Swedish case

Currently, the immigrant population in Sweden constitutes around sixteen percent of the total number of inhabitants (Statistics Sweden 2015). Many immigrants come from countries in the Middle East, Africa, South-East Asia and Eastern Europe, and in terms of educational qualifications and labor market participation they constitute a heterogeneous group.

Sweden combines a recent history of high immigration rates and subsequent ethnic segregation with relatively equal living conditions for all inhabitants; a result of the strong Swedish welfare state institutions. One of the major goals of Swedish educational policy has been to decrease the association between social background and educational attainment, to facilitate equality of opportunity (Björklund et al. 2005; Erikson and Jonsson 1996). These efforts might imply that levels of ethnic disadvantage due to either social background or school segregation would be rather low in Sweden.

Nevertheless, the last wave of the PISA survey demonstrates that Swedish school results (particularly in math) are more differentiated on the basis of native/immigrant background than is the case in many other countries (Skolverket 2013). Differences in grades between Swedish-born and immigrant students have increased over the last two decades, at the same time as the between-school-variation in students' achievement has increased (Skolverket 2012). Over the same period, levels of segregation have increased in Swedish schools, in

terms of both parental education and immigrant background, whereas there is more of a fluctuation in levels of segregation based on parental income. It should be noted, however, that despite increasing levels of segregation on several dimensions, when viewed in an international perspective, the level of school segregation in Sweden is somewhere in the lower to middle range (Böhlmark, Holmlund and Lindahl 2015). Nonetheless, increasing levels of segregation, together with increasing differences in educational outcomes by family background, indicate that Sweden may be no different from the rest of the western world when it comes to school segregation and its effects on students' educational outcomes.

DATA AND VARIABLES

The data set on which the empirical analyses in this paper are based includes information on all students (1,574,237 individuals) who finished the ninth grade in all lower secondary schools between 1998 and 2012 (1,339 schools). All data are drawn from a compilation of Swedish population register data located on Statistics Sweden's servers, accessed via the Institute for Analytical Sociology at Linköping University. The information on individual students comes from an official school register of ninth-graders (*Åk9-registret*), which includes individual-level information on grades and other school outcomes as well as a unique school identifier. The school identifier makes it possible to link students who attended the same school in the ninth grade. The data set was matched (using unique personal identification numbers) with register data on parents, which was for the most part drawn from tax registers and registers of education, civil status and the like. Thus, the data on the students' school results are combined with information about characteristics of their families, as well as information on their schoolmates' families.

We excluded very small schools with a total of fewer than 20 ninth grade students in the year in question and schools where the total number of ninth grade students for the whole

observation period (1998-2012) was less than 300 (85,416 students, or 5 percent of the initial population). Small schools include a comparatively large proportion of schools for children with special needs and religious schools with a restricted intake and special curricula, which means that these schools have a student composition that is distinctly different from that of ordinary public and independent schools.

Outcome variables

Almost all Swedish children pass through the publicly financed lower secondary school system between the ages of 7 and 16. There is little streaming or tracking in lower secondary school. Teachers grade students in 16 different subjects on a four point grading scale: no pass, pass, pass with distinction, and pass with special distinction. For applications to upper secondary school, these grades are recoded to 0, 10, 15, and 20 points, respectively.

According to national regulations, grades measure students' knowledge as evaluated in relation to pre-determined goals in a national curriculum. Each school must conduct identical national tests in core subjects (Swedish, Math and English) and the National Agency of Education can check to make sure average school grades do not deviate too much from the school's average results on these tests. Teachers are also encouraged to "use all kinds of information" to assess students' knowledge, including classroom performance. Thus the grades reflect both the students' performance on general ability tests and the teachers' subjective impressions of the students. The teacher-assigned grades have an impact on the range of options available to students when they apply for a place at upper secondary school (age 17-19).

The study focuses on two educational outcomes: (1) *Grades* are measured by the summarized score of the grades in the 16 highest graded subjects, varying between 0 ('no pass' in all subjects) and 320 ('pass with special distinction' in all subjects). These are the grades the students use to apply for secondary school. (2) *Eligibility for secondary school* is measured as

whether a student has a ‘pass’ in English, Math, and Swedish (for immigrants: Swedish as a second language).

Students and their families

The students’ immigrant status distinguishes between (1) *first generation immigrants*, defined as all students born abroad to two foreign-born parents, and (2) *second generation*, indicating that the student was born in Sweden but both his/her parents were born abroad. Individuals with parents of mixed Swedish/non-Swedish background are defined as *native students of Swedish background*. Adopted children who were born abroad are classified as native children. These children usually arrive in Sweden at a very young age. In all analyses, we perform separate regressions for these three groups. .

Register data do not include information on ‘ethnicity’. We therefore operationalize ‘ethnicity’ in terms of the individual’s country of birth. Note that persons with the same country of birth may belong to different ethnic groups. For example, individuals who were born in Turkey may be ethnic Turks, Kurds, Assyrians, Roma, Arabs, Greeks, or Armenians. For some of the smaller immigrant groups, we use *region* of birth rather than *country* of birth. This is due to a regulation that is employed by Statistics Sweden in order to protect anonymity. Appendix A includes a specification of the countries and regions of birth that we can distinguish in our data.

The parents’ *level of education* is based on the highest level of education achieved by any of the parents. We distinguish here between lower secondary school, upper secondary school, tertiary education of less than two years, and tertiary education of two years or longer. The available information on immigrants’ education is for the most part drawn from a survey that is conducted among all immigrants. We also include a category for those with missing information in relation to parental education (4.6 percent of the immigrant population and 0.1 percent of the native population).

Three variables measure family demography: *living in an intact family* (in the same household as both of the biological or adoptive parents), *number of siblings*, and parity: *the position of the student in the sibling set*.

Parental employment is coded separately for each parent and is defined as having annual earnings above 70,000 SEK (in 2012 prices). This allows us to capture the effect of being brought up in a family with zero, one, or two working parents.

Household income is measured as the sum of the disposable incomes of both parents, averaged over the three most recent years. We include income from employment, business, and taxable cash benefits, in 2012 prices. If information is missing for any of the years, we use the non-missing years to construct the measure. If the child is only registered to one parent, or if one of the parents is deceased or living abroad, the income from the existing parent is used as household income. A lack of information for one of the parents is particularly common among immigrants, where we have no information about the father's income for 24 percent of the students. The household income variable is included in the analyses as yearly constructed quintiles, with a separate category for the (few) students for whom we have no information on either the mother or the father's income (0.1 percent of the Swedish students and 0.8 percent of the immigrant students)

Finally, for the immigrant population, we include a variable measuring age at immigration, distinguishing between those who arrived prior to age 7, at age 7-9, age 10-12, and age 13 and those who were older than this at the time of their arrival. All variables are, if nothing else is specified, measured during the year the student completed the ninth grade (usually at age 16).

School context

All of the contextual variables are constructed in the form of jack-knives and exclude the individual's own contribution to the measure. They thus reflect the composition of 'the

others' within the school. Our critical independent variable is *the proportion of (first generation) immigrant schoolmates* in the school's ninth grade. We also study *the proportion of immigrant schoolmates who arrived at age 10 or older*. This variable may be crucial to understanding the problems that may arise in the presence of newly arrived students with special needs, such as a need for extra language training or help with emotional problems, which will demand special attention from teachers, and which may lower the quality of classroom instruction. In order to further examine the importance of these kinds of special needs we use the students' country of birth to create a measure of *the proportion of immigrant schoolmates born in countries/regions from where the largest stream of immigrants to Sweden consists of refugees*. These include immigrants from Afghanistan, former Yugoslavia, Eritrea, Ethiopia, Iraq, Iran, Lebanon, the Middle East, Somalia and Syria (62 percent of the total immigrant population in our data set).

Besides our measure of immigrant density, we can also gauge the proportion of co-ethnics at school, i.e., individuals sharing the same country of origin. We construct two variables: *the proportion of co-ethnics* in the ninth grade, and *the proportion of non-co-ethnic immigrants*. Both variables are based on combining the country (or in some cases the region) of birth of the student with the country of birth of the other students in the ninth grade in a given school and year. In all analyses, we include a measure of *the proportion of schoolmates who have at least one parent with two years or more of tertiary education* in order to adjust for the socioeconomic composition of schools.¹

Finally, we include a school-level variable measuring whether the school is an independent or public school. The proportion of pupils attending independent schools increased from 1.5 percent in 1998 to 9.8 percent in 2012, due to a policy change in 1992, making this an important control variable.

ANALYTICAL STRATEGY

Identifying social-interaction effects

Analyzing social environment effects on individual outcomes involves several methodological problems (Hanushek, Kain and Rivkin 2009; Hanushek et al. 2003; Dietz 2002; Manski 2000). As noted, variation in individual outcomes between, for instance, different schools can be the result of three different types of processes: environmental effects, selection effects, and social interaction effects. In this study we are particularly interested in the social-interaction effect, and a crucial task is to disentangle it from any environmental or selection effects. That is, we want to assess whether attending an ethnically segregated school affect educational outcomes *net of the overall school quality and net of selection processes making students attend more or less segregated schools.*

In the case of school segregation and educational outcomes, the overall school quality constitutes an environmental effect. For instance, if the concentration of immigrant students in a school is negatively correlated with school quality, resource allocation and teacher recruitment (e.g., Clotfelter, Ladd and Vigdor 2005; Hanushek, Kain and Rivkin 2004), which in turn affect school results, we have an example of environmental effect which may bias our results, unless properly adjusted for.

Furthermore, it is assumed that *selection* on observed and unobserved characteristics causes an upward bias in estimates of environmental effects, because unobserved (omitted) variables (such as parental aspirations), on the basis of which families sort themselves into residential areas and schools, are related to the studied outcome, i.e. children's school success. This is often called the endogeneity problem. Because families normally live in areas according to their resources, children from affluent families have a high probability of ending up in affluent social environments, often with few immigrant schoolmates. These children are also likely to do well in school, as a result of their family background, which risks exaggerating

the impact of school segregation on school outcomes. In order to capture the effects of the social environment it is therefore necessary to net out the potential effects of selection based on family characteristics.

No entirely satisfactory solution exists for such identification problems in observational studies such as ours. Thus, causal interpretations of the results must remain tentative.

However, we do think that the issue can be addressed in ways that give high credibility to a causal interpretation of the results. In our study we deal with selection bias in several ways:

First, we control for a set of empirically crucial and well-measured variables on both individual background and the socioeconomic composition of schools.

Second, we address the problem of unobserved school heterogeneity (i.e. differences in quality between schools with more or less immigrants) by applying what is generally known as a school fixed effects model. The use of school fixed effects constitutes an efficient way of controlling for all of the characteristics of schools, and of the students attending a given school, which do not vary within the school over time (such as unmeasured resources and teacher quality).

Third, we estimate family fixed effects models to control for unobservable family characteristics that are constant across siblings and time and that can influence students' grades and school choice (Hanushek, Kain and Rivkin 2002; Plotnick and Hoffman 1999).

Under the assumption that parents' resources and aspirations are the same for all of their children (or that they at least do not systematically co-vary with their children's school characteristics), we adjust for all such characteristics. These models use the sibling set as the unit of observation, excluding families with only one child. The identification strategy of this model hinges upon variation in the contextual variables within a sibling pair. There are two sources for variation in the school context: siblings are born some years apart and the school's characteristics change over time, or siblings attend different schools. Accordingly, we

estimate separate models for *stayers* (where all siblings attend the same school) and for *movers* (where at least one sibling attends a different school) in order to scrutinize which type of identification we are dealing with. We believe that the family fixed effects method is a very efficient way of addressing the endogeneity problem. This strategy also implies that two siblings who attend the same school in the same year (e.g., twins) do not contribute to the estimation. (See Appendix B for a description of the within-family variation of our independent and dependent variables.)

Fourth, we estimate a two-way fixed effects model of families and schools, using a Stata algorithm developed by Cornelissen (2008). In this specification, both the family and the school are assumed to have time-invariant characteristics that influence educational success. This idea has a lineage from labor economics, where it has been used with the aim of capturing both firm and worker heterogeneity (Abowd and Kramarz 1999). The model is identified by sets of siblings that attend at least two different schools (*movers*) in order to provide the variance needed to identify the two fixed effects. The motivation behind this model is that parents may react to poor school performance by the older sibling and choose another school for the younger sibling. The siblings' schools may differ in the proportion of students comprised of immigrants, but also in overall quality, as a result of segregation-generated selection processes in relation to schools. Thus if we do not adjust for differences in overall school quality, by means of a two-way fixed effects approach, the results from the family fixed effects models may indicate a spurious relationship between the proportion of immigrants and school performance, when in fact it is the overall quality of the schools that differs between the siblings.

It should be noted that reducing upward bias by using an extensive set of control variables or different fixed effects models may under some circumstances introduce a downward bias and produce conservative estimates. This is because the control variables are endogenous to the

extent that ethnic segregation impacts on our indicators of parental and household resources (cf. Duncan, Connell and Klebanov 1997; Sampson, Morenoff, and Gannon-Rowley 2002). For example, if residential segregation means that immigrants have smaller chances of getting a job and avoiding poverty (as a result of mechanisms outlined by Wilson 1987, for example), then controlling for (parental and school-level) income and unemployment will lead to a conservative estimate of the influence of ethnic segregation. It is also likely that our school fixed effects model partials out differences between schools that are affected by consequences of ethnic segregation (e.g., if the most qualified teachers avoid certain schools because they have a high concentration of immigrant students). The family fixed effects method, while efficiently addressing endogeneity problems, is a very demanding test due to the fact that changes in school context between siblings (both stayers and movers) might be rather small. Small changes in the school context allow model identification, yet one cannot expect that such limited variation in an independent variable will produce large effects in the dependent variable. This issue is even more amplified for the two-way school and family fixed effects models.

RESULTS

Descriptive results

Table 1 presents descriptive statistics for school-level variables separately for students of Swedish background, for the Sweden-born children of immigrants (second generation) and for foreign-born students (first generation immigrants). (Appendix C presents descriptive statistics for the full set of variables.)

[TABLE 1 ABOUT HERE]

On average, immigrants account for 5 percent of the schoolmates of students of native Swedish background, and newly arrived immigrants account for only 2 percent of this group's

schoolmates. By contrast, an average of 18 percent of the classmates of first generation immigrant students were born abroad, and 8 percent were recently arrived immigrants. On average, *co-ethnic* students account for 3 percent of the classmates of immigrant students, and there are first generation immigrants attending schools at which as many as 42 percent of their schoolmates are co-ethnics, that is, foreign-born students from the same country.

Furthermore, there are clear differences in levels of school achievement. Whereas 91 percent of the native students are eligible for upper secondary school upon finishing the ninth grade, this is the case for only 84 percent of the second generation students, and as few as 68 percent of the first generation immigrants. These differences are also visible in the average grade point scores of the three student groups (208, 195 and 169, respectively).

Multivariate results

Tables 2 through 4 present results from analyses that estimate the effect of different measures of ethnic concentration in school on average grades using OLS regression analyses (Table 2), school fixed effects models (Table 3), family fixed effects models and two-way fixed effects models (Table 4). This is followed by a presentation of similar sets of analyses focused on eligibility for upper secondary education (Tables 5, 6 and 7). Finally, for immigrant students only, we present analyses of how the structure of the ethnic composition in school matters for their educational outcomes, by studying the importance of the proportion co-ethnics in school (Table 8).

Average grades – OLS analyses

We start by presenting results from OLS regressions estimating the effects of ethnic school segregation on students' average grades (Table 2). We adjust for the non-independence of observations within the same school by clustering the standard errors on the school level. We include three variables to measure ethnic school segregation; a) the proportion of schoolmates

born outside Sweden; b) the proportion of schoolmates who arrived in Sweden at age 10 or later; and c) the proportion of schoolmates who were born in typical refugee countries. For each of these three independent variables, we estimate (1) a model only controlling for the proportion of children with tertiary educated parents, calendar year, and an indicator of whether the schools is an independent school and (2) a model also controlling for a range of family characteristics. Models 1 and 2 include the native Swedish students, Models 3–4 include students of immigrant background who were born in Sweden (2nd generation), and, Models 5–6 include students of immigrant background born abroad (1st generation).

[TABLE 2 ABOUT HERE]

We start with students of Swedish background. Model 1 shows that grade scores are on average 37 points lower in schools in which all students are immigrants than in schools with no immigrant children. To understand the size of this effect, we compare the grades of a Swedish student with a median share of immigrant classmates to those of a Swedish student who attends a school with an immigrants' median share of immigrant classmates. The median share of immigrant classmates for a Swedish student is 3 percent, compared to 13 percent for immigrant students (Table 1), that is, a 10 percentage point difference. This corresponds to a reduction in grade scores of around 3.7 points for the student exposed to more immigrants (the estimate is obtained by multiplying 37.5 with 0.10). Thus the association is rather weak. The grade scale ranges from 0 to 320 points, where 3.7 points corresponds to less than an increase from pass to pass with distinction in one (out of sixteen) subjects.

The associations between the two other measures of immigrant concentration and grades are stronger. In particular, the corresponding coefficient for the proportion of immigrant children who arrived in Sweden at the age of ten or later is large, at 82 points.² However, for a ten percentage point difference this still only corresponds to a reduction in grades of 8 points.

In Model 2 we analyze (for students of Swedish origin) whether the effects of ethnic segregation in schools are due to the selection of students from poor social conditions to schools with a high proportion of students of immigrant background. We do this by adding controls for family resources. For all of our measures of ethnic segregation, the control variables introduced in Model 2 (and later in Model 4 and Model 6) reduce the strength of the association found in Model 1. There is some variation in the size of this reduction depending on the definition of ethnic segregation employed. In general, the reduction is not very substantial with the exception of that associated with the proportion of newly arrived immigrants, where the control variables are more important. Thus differences in grade scores between schools of varying ethnic composition are in part due to a selection of students from families with different levels of resources (as measured here). However, a large part of the between-school variance remains unexplained. The reduction is most noticeable for our measure of socio-economic segregation (proportion of students whose parents have an academic education).

Models 3 and 4 present the results for the second generation immigrants. The raw coefficients associated with ethnic segregation in Model 3 vary with the definition of segregation, but are generally in line with the coefficients for the students of Swedish background. Conditioning our analysis on family characteristics (Model 4) reduces the strength of the association between our indicators of ethnic school segregation and grades. The reduction is often greater in the case of the second generation immigrants (as compared with the models for students of Swedish background) indicating that social selection is stronger for the second generation students.

Finally, in Models 5 and 6 we analyze first generation immigrant students. As can be seen from Model 5, the unconditional estimates of the associations between all of our indicators of ethnic school segregation and grades are much stronger for the first generation immigrants

than for the other two student groups. For immigrant students, a 10 percentage point increase in the share of immigrant classmates is associated with a reduction of 6.5 grade points (-66*.10). Applying controls reduces the estimates substantially. However, for all of the measures of ethnic school segregation presented in Table 2, the conditional associations in Model 6 are stronger than those shown in Models 2 and 4. Thus so far the results indicate that ethnic school segregation seems, above all, to be a problem for first generation immigrants. Another interesting result is that school segregation linked to parents' educational resources (the proportion of students of academic background) seems to be most important for students of Swedish origin and for the second generation students.

Average grades – school fixed effects

In the next stage of our analyses (Table 3), we address the problem of unobserved school heterogeneity. The ethnic segregation effect shown in the section above may be partly related to differences in school quality between schools with a larger or smaller proportion of immigrants. We replicate the analysis of school grades using a set of school fixed effects models (school dummies). In doing so, we control for all characteristics of the school and of the students that do not vary within the school across years (such as unmeasured school resources, teacher quality etc.). All of the individual- and school-level control variables included in Table 2 are also included in Table 3.

[TABLE 3 ABOUT HERE]

In Model 1 we analyze how shifts in the ethnic composition *within* schools affect the grades of native Swedish students attending these schools at different years during the period studied. The estimates are negative, and very similar for all measures of ethnic school segregation, indicating that an increase in the proportion of immigrant students over time results in a deterioration in grades. However, the estimates also indicate that the change in grades is rather minor. Experiencing a 50 percent increase in the proportion of immigrant students,

which is a large *within-school* increase, would on average lead to deterioration in students' results of around 7 points on a scale that ranges from 0 to 320.

Interestingly, for the second generation immigrant students (Model 2) the results indicate that increasing (or decreasing) the degree of ethnic school segregation does not affect students' grades. Finally, for the first generation students (Model 3) the results are mixed. The effects of changes in the general proportion of immigrant students, and in the proportion of students from refugee countries, are minor. For the proportion of students who arrived in Sweden at the age of 10 or later, the effect is similar to the effect shown in Model 1 for native Swedish students. Thus, in models which are not biased due to omitted school-level characteristics, the effects of ethnic school segregation on grades are generally minor. This effect is estimated with relatively low precision. The standard error indicates a high degree of heterogeneity between schools that experience a shift in the proportion of immigrants. It should also be noted that in all these models the estimate for the proportion of students from an academic background is also of minor importance.

Average grades – family fixed effects and two-way fixed effects

School fixed effects models constitute an efficient instrument for controlling for all school and student characteristics that are constant over time. However, if the characteristics of students (and their families) change during the period in ways that we cannot observe, our results can be biased due to unobserved heterogeneity related to students' families. For instance, if the share immigrant pupils increase in a school over time and this is accompanied by the native pupils with the best grades leaving the school, a school-fixed-effects model will overestimate the impact from the share of immigrants on the grades of native pupils.

One solution to this problem is to estimate family fixed effects models which allow us to control for unobservable family characteristics that are constant across siblings and time. We do this in several steps. First, we estimate family fixed effects models separating 'stayers' (i.e.

siblings who attend the same school) from ‘movers’ (i.e. siblings who attend different schools). Second, we estimate two-way fixed effects models using both family and school identification. This last procedure amounts to simultaneously applying a set of dummies for schools and families. This means that we adjust for all time-invariant family and school characteristics, such as overall support from parents and school quality. In all models we include the same control variables as in Tables 2 and 3, such as parental employment, separation or divorce, number of siblings, *etc.*

In these analyses we merge first and second generation immigrants. Since we are analyzing siblings within families, one sibling may have been born prior to arrival to Sweden while other(s) may have been born in Sweden; thus some families may comprise children categorized as both first and second generation students. We control for whether the sibling is foreign born, and, if so, their age at immigration.

[TABLE 4 ABOUT HERE]

Starting with families of Swedish origin, Table 4, Model 1, shows the results for siblings who attend the same school (‘stayers’). For stayers, the effects of within-family variation in ethnic school segregation are very small. For movers (Model 2), we find a considerably stronger pattern. The within-family variation in ethnic school composition indicates that the sibling who was exposed to a larger proportion of immigrant schoolmates ends up with lower grades. The within-family estimate is particularly strong for those siblings who experience large differences in the proportion of immigrant students who arrived in Sweden at the age of 10 or later.³

Model 3 presents the estimates for our preferred model, which combines school and family fixed effects, and thus simultaneously adjusts for possible differences in school quality and unmeasured family characteristics. It should be noted here that the identification assumption for the two-way fixed effects model is that sibling pairs do not attend the same school

(Andrews, Schank and Upward 2008). Thus our model is identified for movers only. All of the estimates indicating effects of school segregation on grades (Model 3) are close to zero and not significant. Thus, for students of Swedish origin, taking into account both quality of schools and ‘quality’ of families (in terms of dummy variables in two-way fixed effects models) produces the finding that the share of immigrants has no effect on their final grades from lower secondary school.

In the next three models we analyze students of immigrant background. For stayers (Model 4), the within-family fixed effects analysis indicates a relatively stable, but weak, pattern showing that an increasing proportion of foreign-born students at school results in *higher* grades for the sibling exposed. For movers (Model 5), the result is just the opposite; the sibling with more immigrant schoolmates tends to have lower grades. Again, given the features of the variable measuring grades, which ranges from 0 to 320, the substantive importance of the results of Models 4 and 5 should not be overstated. The final test of the effect of ethnic school segregation on grades by means of a two-way fixed effects model (Model 6) indicates that the effect of ethnic segregation on grades for students of immigrant background is not significant and rather weak (all coefficients below 7), when adjusting for overall school and ‘family’ quality. In the unrealistic scenario of the first sibling attending a school with no immigrant students and the second sibling attending a school where 100 percent of students were immigrants, the only “gain” in grades would correspond to an increase from pass to pass with distinction in one of 16 subjects.

Eligibility - Linear Probability Models

Table 5 presents results from analyses of eligibility for upper secondary education. The dependent variable is in this case binary and we use linear probability models (LPM) to estimate the effects of the ethnic composition of schools on the probability for being eligible for secondary school. LPM are equivalent to OLS regressions with robust standard errors. It

has been customary in the social sciences to use, for instance, Poisson or logistic regression when estimating models with binary dependent variables. Yet these methods involve problematic assumptions about the error distributions and provide coefficients that cannot be compared across models and specifications (Winship and Mare 1984; Mood 2009), which is why we rather use LPM in the present study. Angrist and Pischke (2009) show that LPM is a valid estimator in cases like ours.

We follow the same analytical steps as we used when studying grades.

[TABLE 5 ABOUT HERE]

Starting with students of Swedish background, Model 1 shows the association between eligibility and different measures of ethnic school segregation and the proportion of children of academic background. The raw estimate of -0.24 in the first row reflects a very unlikely 100 percentage point change in the share of immigrants. A more plausible change of 10 percentage points would imply that the probability of being eligible for upper secondary education decrease by around 2.5 percentage points ($-0.24 \cdot 0.10$). The result is similar when the focus is directed at the proportion of schoolmates from typical refugee countries. The corresponding coefficient for the proportion of newly arrived students is considerably higher (-0.48).

In Model 2 we introduce control variables to account for the selection of students from poor social conditions to the schools attended by students of immigrant background. The association between our measures of ethnic segregation and eligibility for upper secondary school are somewhat reduced, but the remaining part still indicates that for students of Swedish background there is a large between-school variation in levels of eligibility that is linked to the share of immigrants in the school.

In Models 3 and 4 we analyze native-born students of immigrant background. The raw association in Model 3 between segregation and eligibility is strong irrespective of which

indicator of segregation is employed. The introduction of controls (Model 4) reduces the strength of the association but it remains substantial. Finally, turning to the immigrant students born outside Sweden (Models 5 and 6), we find that that their probability of being eligible for upper secondary education is much lower in schools with a large proportion of immigrant students. The differences are rather large. For instance, in the more restrictive model (Model 6), the probability of being eligible decreases by around 3.5 percentage points if the proportion of immigrant children increases by 10 percentage points (-0.34×0.10). For the alternative measures of ethnic school segregation, the differences are much larger; in particular we find very large negative associations related to the proportion of immigrant students that arrived in Sweden after the age of 10. Thus far, then, ethnic segregation in schools appears to have a powerful effect on the opportunities for future educational careers among young people in Sweden.

Eligibility - Linear Probability Models with school fixed effects

Turning to the analyses of school fixed effects (Table 6), we find that the effects of school segregation on eligibility are much reduced for all of the categories of students analyzed.

[TABLE 6 ABOUT HERE]

For students of Swedish origin (Model 1) and first generation immigrants (Model 3), the results are similar and the effects of ethnic segregation are more or less the same irrespective of which measure of segregation is employed. Parameter estimates ranging from -0.09 to -0.13 indicate that the ethnic composition of schools has to change rather radically to produce substantial outcomes. However, during the period analyzed such a change has in fact occurred in some Swedish schools; for example, 11 percent of schools have experienced an increase of at least 10 percentage points in the proportion of immigrant students. Thus, the results indicating an effect of ethnic segregation on eligibility may be of substantial interest. For the

second generation (Model 2), however, the effects of segregation are rather weak. The same applies for our measure of the educational background of students.

Eligibility – Linear Probability Models with family fixed effects and two-way fixed effects

In Table 7 the analyses from Table 4 are repeated with a focus on eligibility for upper secondary education.

[TABLE 7 ABOUT HERE]

Our models for siblings from native Swedish families (Models 1, 2, and 3) indicate that for the probability of being eligible to secondary school, there is a negative effect of ethnic segregation. A within-family increase in the share of immigrants produces a decrease in the probability of being eligible. For stayers (Model 1), the effects are rather weak. However, for movers (Model 2) the effects are larger, ranging from -0.11 to -0.29. Introducing the two-way fixed effects model (Model 3) reduces the strength of the effects, and the remaining effect of the proportion of newly arrived immigrants in a school (-0.12) must be interpreted as minor. A within-family movement from a school where 10 percent of the students are relatively newly arrived immigrants to a school with no newly arrived immigrants results in a one percentage point increase in siblings' chances of being eligible for upper secondary school, net of differences in overall school quality.

In the final three models we focus on immigrants and the children of immigrants. For stayers (Model 4), the effects are close to zero. For movers (Model 5), however, the within-family variation is greater (with estimates ranging from -0.11 to -0.19). If one of the siblings is exposed to a school with a large share of immigrants, the likelihood of being eligible for upper secondary school is lower for this individual than for a sibling who experiences a school environment with a lower degree of segregation. The final model shows that a substantial part of the association between ethnic segregation and eligibility remains even when the two-way fixed effects models are employed. However, even for the largest estimates

– that is, those associated with the proportion of immigrant students and of students from typical refugee countries (-0.08 and -0.09) – a 10 percentage point increase in the share of immigrants is still associated with less than a one percentage point lower likelihood of eligibility.

Thus our analyses of eligibility for upper secondary education indicate that ethnic segregation in schools is of some importance for the future educational careers of young Swedish students, although the effect size is minor. As was shown in the analyses of grades, for the majority of students who are relatively successful in their studies and who complete their compulsory schooling with adequate grades, ethnic school segregation is of minor or no importance. Among weaker students, however, who are on the borderline of qualification for the regular programs in upper secondary school, ethnic school segregation may impair their levels of school achievement.

In order to gauge whether these results hold also for the students not included in this final set of analyses (that is, for those without siblings or where all siblings attend the same school) we have repeated our initial OLS analyses from Table 5 including only the individuals included in the two-way fixed effects models. The coefficients are similar to the OLS-regressions with control variables presented in Table 5, Models 2, 4, and 6. This indicates that it is not the restricted sample that creates these effects.

Structure of ethnic school segregation

In addition to the general level of the concentration of immigrant students, the nature of the ethnic composition at school may also affect students' educational outcomes. For students of immigrant background a high proportion of immigrant students from many different countries may be detrimental while a high proportion of co-ethnics (or co-nationals) may be beneficial. In this section, we disaggregate the fraction of immigrant students in two parts: the proportion

of immigrants from other countries/regions and the proportion of immigrants from the same country/region (see Appendix A for a specification of countries and regions).

[TABLE 8 ABOUT HERE]

In the analyses presented in Table 8 we exclude all students who were born in Sweden, that is, we only include first generation immigrants. We begin with an OLS-analysis that includes family controls and a control for country/region of birth. In the first panel of Table 8 we present the results for grade point scores and in the second (lower) panel the results for eligibility. As can be seen from Model 1, segregation on both of the dimensions is negative for both of the outcomes studied. In this model, the proportion of co-ethnics is more strongly (and negatively) associated with school results than the proportion of non-co-ethnics. One of the reasons may be that strong ethnic communities reside in particularly disadvantaged areas with schools of relatively poor quality. To understand the potential selection processes generating these results, we apply a school fixed effects model (Model 2). The outcomes differ substantially from the results described above. The proportion of non-co-ethnic immigrants is still negatively (but marginally) related to grade point scores. As regards the proportion of co-ethnics, however, the coefficient for grade point scores is close to zero and is not significant. For eligibility, there is still a negative effect associated with the proportion of non-co-ethnics in school, but the effect of co-ethnic concentration is smaller (and not significant).

Turning to the family fixed effects models (reported in Models 3 and 4), the estimates differ substantially from those found in the Models 1 and 2.

For ‘stayers’ (Model 3), there remain weak negative effects of non-co-ethnic immigrant schoolmates on grade point scores and on the probability to be eligible to secondary school. However, for the proportion of co-ethnics, the estimates change in both direction and magnitude. For both grades and eligibility, a within-family increase in school results is

positively related to the proportion of co-ethnics at school, meaning that the sibling who has more schoolmates from the same country/region of birth systematically achieves better educational outcomes than the sibling who attended a school with fewer co-ethnics. When interpreting the substantive meaning of the coefficients, it is important to bear in mind what effect a realistic within-family change in the ethnic composition at school might produce. A change of 10 percentage points in the proportion of co-ethnic students between siblings attending the same school in different years produces an estimated difference of 5.5 grade points ($55 \times .1$), or a 4.4 percentage point difference in probability to be eligible to secondary school ($.44 \times .1$). 5.5 grade points is equivalent of the difference between pass and pass by distinction in one subject, and can hence not be considered to be a very large difference, by contrast to what an interpretation of the raw estimate might lead one to expect. For eligibility, a difference of 4.4 percentage points is less trivial, as the overall probability for being non-eligible among immigrants is 32 percent.

For ‘movers’ (Model 4), the results are different. Here the negative estimates associated with the proportion of non-co-ethnic immigrants are clearly stronger for both of the outcomes analyzed. The influence of the proportion of co-ethnics is slightly weaker in the analysis focused on grade point scores, and is close to zero in the analysis focused on eligibility.

Finally, we report the results from our preferred model (Model 5), in which we account for the quality of schools and families simultaneously, by means of a two-way fixed effects model. These estimates are not subject to selection processes which might bias analyses of how the ethnic environment of schools influences school results. As has already been discussed, these models can only be estimated for movers. The identification assumption is that families have at least two children and that these children attend at least two different schools, with varying degrees of ethnic segregation.

The effect of the proportion of non-co-ethnic immigrants is not significant or strong for either grade points or eligibility after adjusting for both family and school quality simultaneously (Model 5). However, the proportion of co-ethnics still has a positive and significant impact on grade scores, although the magnitude of the effect is weak (similar to that found for ‘stayers’ in Model 3). As regards eligibility, the coefficient describing the effect from co-ethnics is weaker than in Model 3, although it still is of non-trivial magnitude. Applied to a realistic scenario of a change of 10 percentage points in the proportion of co-ethnic schoolmates, the sibling exposed to more co-ethnics will have a 1.5 percentage point higher likelihood of being eligible for upper secondary school ($.15 \cdot .1$), on a baseline where 68 percentage of all pupils are eligible. However, the measure is subject to a lot of noise and does not reach statistical significance.

In this section, we have distinguished two dimensions of ethnic school segregation and two educational outcomes. Having a large number of immigrant peers from *different* countries/regions of origin seems to have only negligible effects on the school results of first generation immigrant students. However, having a large number of immigrant peers from the *same* country of origin has a *positive*, albeit rather weak, effect on the educational outcomes of immigrant students.

DISCUSSION AND CONCLUDING REMARKS

The academic and social policy relevance of the question of whether ethnic segregation reduces children’s educational opportunities is obvious. What is far from obvious, however, is how such segregation effects should be estimated empirically. Much earlier research has shown that there is a negative association between the average performance of students and the proportion of immigrants or children of immigrants in a school. Interpreting such results would be straightforward given an assumption that the distribution of immigrant students across schools is a random process and that resources are evenly distributed between schools.

Under such assumptions, the negative association between educational progress and the ethnic composition of students could be interpreted as a negative peer effect.

However, in the real world, the nature of segregation processes does not function according to these assumptions. In this paper, we have tried to disentangle the effects of three different types of social processes which are likely to affect students' educational outcomes: selection effects, environmental effects, and social interaction effects. The aim of the present study has been to identify causal social interaction effects of the concentration of immigrants in schools on students' educational outcomes at the end of lower secondary school.

There are strong theoretical reasons to expect that a high fraction of immigrant students may have negative consequences for the learning environment in a classroom. A lack of proficiency in the language of instruction, a high prevalence of students with traumatic (their own or their families') experiences, and many examples of low returns to education in the parental generation might lead to a deterioration in opportunities and in the willingness to learn among students. On the other hand, if the ethnic environment in schools is characterized by a high prevalence of students from the same countries of origin, the shared ethnic identity might constitute a significant resource. Close social relationships in an ethnic group may generate social capital that positively contributes to students' school involvement and educational outcomes.

In the empirical part of this paper we analyze an exhaustive Swedish data set which includes all 1.5 million students who left lower secondary schools during the period 1998-2012. We address two important educational outcomes at the conclusion of lower secondary school: students' grade point scores and whether or not they are eligible for upper secondary school. Our analytical strategy for identifying effects of social interaction is to combine school fixed effects, family fixed effects and two-way fixed effects. This design, we believe, provides

sound analytical models which can be interpreted in terms of the causal effects of schools' ethnic composition on students' educational outcomes.

The results are mixed. For the most part, the estimated effects of ethnic segregation are close to zero. Ethnic segregation in Swedish schools, *per se*, does not have any substantial effects on the grade point scores of students. The differences between schools noted in a standard OLS analysis seem primarily to be an artifact of selection and of differences in school- and teacher quality. When we adjust for this, by shifting the focus to the impact of changes in the percentage of immigrants (within schools or between siblings), the estimated differences disappear.

For the weakest students, who were on the borderline for eligibility for upper secondary school, we found a rather small negative impact of attending a school with a large proportion of immigrant schoolmates. This negative effect was found for both native Swedish students and first generation immigrant students. Thus, if ethnic school segregation matters, it mainly has (weak) effects on students who are at the lower end of the achievement scale. A corresponding finding was identified for the effect of the fraction of co-ethnics on eligibility. In this case, students very near the bottom of the grade distribution seem to be weakly positively influenced by their (co-ethnic) peer environment.

Even if our analytical model is rigorous and can be used for causal analyses of the impact of different characteristics of the school environment on educational outcomes, there are some limitations that should be mentioned. When parents make decisions about the choice of school for their children, they contribute to schools having a specific social and ethnic composition. This composition may be seen as an opportunity structure in relation to social relations and interactions among the students at school. However, the actual frequency of interactions between students in the same grade and school is unknown. Thus different peer characteristics may only be viewed as proxies for the processes of social influence.

Another limitation (or trade-off), which has already been mentioned, is that while our rigorous methodology may be efficient in bias-reduction, the use of statistically demanding models may simultaneously produce conservative estimates of the processes analyzed. There is no way to completely overcome the range of problems that researchers encounter in trying to arrive at causal conclusions on the basis of observational data. The main advantage with our analytical approach is that we can be fairly certain that any remaining association between school composition and educational outcomes is of a causal nature, which lends our conclusions a high degree of credibility.

The central finding, which may be considered counterintuitive, that ethnic segregation in schools has minor or non-existent (short-term) effects on students' educational outcomes should be seen in the context of the issue of when and how the future educational outcomes of young persons are generated. According to Heckman (2006), educational achievements and careers are based on foundations that are laid down early in life. In the early phases of life, the primary socialization agent is the close family. Family background has a major influence on the formation of ambitions, expectations and norms during the early years of adolescence. This is where the future educational outcomes emerge. The "long arm" of the family of origin seems to trump those processes that revolve around the (ethnic) composition of students in schools.

This is particularly interesting in relation to the widespread phenomenon of white flight behavior (Spaiser et al. 2016), whereby parents go to great lengths to avoid schools with a large proportion of immigrants. Our results indicate that if such parents are mainly basing their behavior on the immigrant composition of the school in question, then their children do not appear to be gaining very much from it.

RESEARCH ETHICS

This research has been approved by the regional ethical review board in Stockholm, Sweden, and by Statistics Sweden's internal equivalent. The researchers have only had access to an anonymized version of the population register data, making it impossible to identify particular individuals. In order to ensure the anonymity of the research population, the data is only accessible through Statistics Sweden's protected servers. Only aggregate data and tables can be exported outside these servers.

REFERENCES

- Åberg, Yvonne, and Peter Hedström. 2011. "Youth unemployment: a self-reinforcing process?" Pp. 201-227 in *Analytical Sociology and Social Mechanisms* edited by Pierre Demeulenaere. Cambridge: Cambridge University Press.
- Abowd, John M., and Francis Kramarz. 1999. "The Analysis of Labor Markets using Matched Employer-Employee Data." Pp. 2629-2710 in *Handbook of Labor Economics* edited by Orley Ashenfelter and David Card. Amsterdam: Elsevier.
- Agirdag, Orhan, Mieke Van Houtte, and Piet Van Avermaet. 2012. «Ethnic School Segregation and Self-Esteem: The Role of Teacher-Pupil Relationships.» *Urban Education*, 47: 1135-1159.
- Akerlof, George. 1997. "Social distance and social decisions." *Econometrica*, 65: 1005-1027.
- Akerlof, George, and Rachel E. Kranton. 2002. "Identity and Schooling: Some Lessons for the Economics of Education." *Journal of Economic Literature*. Vol. XL: 1167-1201.
- Alba, Richard, Jennifer Sloan, and Jessica Sperling. 2011. "The Integration Imperative: The Children of Low-Status Immigrants in the Schools of Wealthy Societies." *Annual Review of Sociology* 37:395-415.
- Allport, Gordon W. 1954. *The Nature of Prejudice*. Cambridge, Massachusetts: Perseus Books.
- Andrews, Martyn J., Len Gill, Thorsten Schank, and Richard Upward. 2008. "High wage workers and low wage firms: negative assortative matching or limited mobility bias?" *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171:673-697.
- Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

- Åslund, Olof, Per-Anders Edin, Peter Fredrikson, and Hans Grönqvist. 2011. "Peers, Neighborhoods, and Immigrant Student Achievement: Evidence from a Placement Policy." *American Economic Journal: Applied Economics* 3: 67-95.
- Bifulco, Robert, Jason M. Fletcher, and Stephen L. Ross. 2011. "The Effect of Classmate Characteristics on Post-secondary Outcomes: Evidence from the Add Health." *American Economic Journal: Economic Policy* 3:25-53.
- Birkelund, Gunn E. and Arne Mastekaasa, eds. 2009. *Integrert? Utdanning og arbeid blant innvandrere og deres etterkommere i Norge*. Oslo: Abstrakt forlag.
- Bitterman, Danuta, and Eva Franzén. 2007. "Residential Segregation." *International Journal of Social Welfare* 16(S1):S127-62.
- Björklund, Anders, Melissa A. Clark, Per-Anders Edin, Peter Fredriksson and Alan Krueger. 2005. *The Market Comes to Education in Sweden*. Russell Sage Foundation.
- Böhlmark, Anders, Helena Holmlund and Mikael Lindahl, 2015. "School choice and segregation: evidence from Sweden." *IFAU working paper 2015:8*.
- Borjas, George. 1995. "Ethnicity, Neighborhoods, and Human Capital Externalities." *American Economic Review*, 85(3), 365-90.
- Brunello, Giorgio and Lorenzo Rocco. 2013. "The effect of immigration on the school performance of natives: Cross country evidence using PISA test scores." *Economics of Education Review* 32:234-246.
- Bygren, Magnus, and Ryszard Szulkin. 2010. "Ethnic Environment during Childhood and the Educational Attainment of Immigrant Children in Sweden." *Social Forces* 88:1305-1329.
- Cebolla-Boado, Héctor. 2007. "Immigrant Concentration in Schools: Peer Pressures in Place?" *European Sociological Review* 23(3):341-56.
- Clotfelter, Charles T., Helen F. Ladd and Jacob Vigdor. 2005. "Who Teaches Whom? Race and the Distribution of Novice Teachers." *Economics of Education Review* 24(4):377-92.

- Coleman, James S., and Thomas Hoffer. 1987. *Public and Private High Schools. The Effect of Communities*. New York: Basic Books.
- Coleman, James S. 1988. "Social Capital and the Creation of Human-Capital." *American Journal of Sociology*, 94: 95-120.
- Cornelissen, T. 2008. "The Stata command `felsdsvreg` to fit a linear model with two high-dimensional fixed effects." *Stata Journal* 8:170-189.
- Crowder, Kyle, and Scott J. South. 2003. "Neighborhood distress and school dropout: the variable significance of community context." *Social Science Research* 32:659-98.
- Cunha, Flavio, and James Heckman. 2007. "The Technology of Skill Formation." *American Economic Review*, 97(2), 31-47.
- Dietz, Robert D. 2002. "The Estimation of Neighborhood Effects in the Social Sciences: An Interdisciplinary Approach." *Social Science Research* 31(4):539-75.
- Duncan, Greg J., James P. Connell, and Pamela K. Klebanov. 1997. "Conceptual and Methodological Issues in Estimating Causal Effects of Neighborhoods and Family Conditions on Individual Development." Pp. 219-50 in *Neighborhood Poverty. Volume 1. Context and Consequences for Children*, edited by Jeanne Brooks-Gunn. New York: Russell Sage Foundation.
- Erikson, Robert. 1994. "Spelar valet av skola någon roll?" Pp. 132-71 in *Sorteringen i skolan*, edited by Robert Erikson and Jan O. Jonsson. Stockholm: Carlssons.
- Erikson, Robert, and Jan O. Jonsson. 1996. "Explaining Class Inequality in Education: The Swedish Test Case." Pp. 1-63 in *Can Education Be Equalized? The Swedish Case in Comparative Perspective*, edited by Robert Erikson and Jan O. Jonsson. Boulder: Westview Press.
- Fekjær, Silje Noack and Gunn E. Birkelund. 2007. "Does the Ethnic Composition of Upper Secondary Schools Influence Educational Achievement and Attainment? A Multilevel Analysis of the Norwegian Case." *European Sociological Review*, Vol. 23 (3): 309-323.

- Fleischmann, Fenella, Karen Phalet, Patrick Deboosere, and Karel Neels. 2012. "Comparing Concepts of Ethnicity in Ethnic Composition Measures: Local Community Contexts and the Educational Attainment of the Second Generation in Belgium". *Journal of Ethnic and Migration Studies* 38(10): 1513-1531.
- Fletcher, Jason. 2010. "Spillover Effects of Inclusion of Classmates with Emotional Problems on Test Scores in Early Elementary School." *Journal of Policy Analysis and Management* 29(1):69-83.
- Fordham, Signithia, and John U. Ogbu. 1986. "Black students' school success: Coping with the "burden of 'acting white'"." *The urban review* 18(3): 176-206.
- Geay, Charlotte, Sandra McNally and Shqiponja Telhaj. 2013. "Non-Native Speakers of English in the Classroom: What Are the Effects on Student Performance?" *Economic Journal* 123(570):F281-F307.
- Gould, Eric D, Victor Lavy and Daniele M. Paserman. 2009. "Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence." *Economic Journal* 119(540):1243-69.
- Hällsten, Martin, Ryszard Szulkin and Jerzy Sarnecki. 2013. "Crime as a Price of Inequality? The Delinquency Gap between Children of Immigrants and Children of Native Swedes." *British Journal of Criminology* 53(3):456-481.
- Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2002. "New evidence about Brown v. Board of Education: The complex effects of school racial composition on achievement." *NBER working paper W8741*.
- Hanushek, Eric A., John F. Kain and Steven G. Rivkin. 2004. "Why Public Schools Lose Teachers." *Journal of Human Resources* 39(2):326-54.
- 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):349-83.

- Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven G. Rivkin. 2003. "Does Peer Ability Affect Student Achievement?" *Journal of Applied Econometrics* 18:527-544.
- Hardoy, Inés and Pål Schøne. 2013. "Does the Clustering of Immigrant Peers Affect the School Performance of Natives?" *Journal of Human Capital* 7(1):1-25.
- Heath, Anthony F., Catherine Rethon and Elina Kilpi. 2008. "The Second Generation in Western Europe: Education, Unemployment, and Occupational Attainment." *Annual Review of Sociology* 34:211-35.
- Heckman, James J. 2006. "Skill formation and the economics of investing in disadvantaged children." *Science*, 312(5782): 1900-1902.
- Hermansen, Are Skeie, and Gunn Elisabeth Birkelund. 2015. "The Impact of Immigrant Classmates on Educational Outcomes." *Social Forces*, 94 (2): 615-646.
- Holmlund, Helena, Josefine Häggblom, Erica Lindahl, Sara Martinson, Anna Sjögren, Ulrika Vikman and Björn Öckert. 2014. "Decentralisering, skolval och fristående skolor: resultat och likvärdighet i svensk skola". *IFAU Report, 2014:25*.
- Jackson, Michelle, Jan O. Jonsson and Frida Rudolphi. 2012. "Ethnic Inequality in Choice-driven Educational Systems: A Longitudinal Study of Performance in Choice in England and Sweden." *Sociology of Education*, 85(2):158-178.
- Jensen, Leif. 2001. "The Demographic Diversity of Immigrants and Their Children." Pp. 21-56 in Rubén G. Rumbaut and Alejandro Portes (eds.), *Ethnicities. Children of Immigrants in America*. Berkeley. Cal.: University of California Press.
- Jensen, Peter and Astrid Würtz Rasmussen. 2011. "The Effect of Immigrant Concentration in Schools on Native and Immigrant Children's Reading and Math Skills." *Economics of Education Review* 30(6):1503-15.
- Jonsson, Jan O. and Frida Rudolphi. 2011. "Weak performance – strong determination. School achievement and educational choice among ethnic minority students in Sweden." *European Sociological Review*: 27:487-508.

- Kao, Grace and Jennifer S. Thompson. 2003. "Racial and Ethnic Stratification in Educational Achievement and Attainment." *Annual Review of Sociology* 29:417-42.
- Lavecchia, Adam M., Heidi Liu and Philip Oreopoulos. 2014. "Behavioral Economics of Education: Progress and Possibilities." *IZA Discussion Papers*, No. 8853
- Lazear, Edward. P. 2001. "Educational production." *Quarterly Journal of Economics* 116:777-803.
- Lee, Jennifer and Min Zhou. 2015. *The Asian American Achievement Paradox*. Russel Sage publication.
- Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies* 60(3):531-42.
- Manski, Charles F. 2000. "Economic Analysis of Social Interactions." *Journal of Economic Perspectives* 14(3):115-36.
- McPherson, Miller, Lynn Smith-Lovin and James M. Cook. 2001. "Birds of a feather: Homophily in Social Networks." *Annual Review of Sociology*, Vol. 27: 415-444.
- Mood, Carina. 2009. "Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It." *European Sociological Review*, 26: 67-82.
- Moody, James. 2001. "Race, School Integration, and Friendship Segregation in America." *American Journal of Sociology*, 107: 679-716.
- Morgan, Stephen L. 2005. "On the Edge of Commitment. Educational Attainment and Race in the United States." Stanford: Stanford University Press.
- Musterd, Sako. 2005. "Social and Ethnic Segregation in Europe: Levels, Causes, and Effects." *Journal of Urban Affairs* 27:331-48.
- Nekby, Lena. 2011. "Comment on Bratsberg, Raaum and Røed: Educating children of immigrants: Closing the gap in Norwegian schools." *Nordic Economic Policy Review. Economics of Education*. 1:253-260.

- Nordin, Martin. 2013. "Immigrant School Segregation in Sweden." *Population Research and Policy Review*. 32(3): 415-435.
- Ohinata, Asako and Jan C. Van Ours. 2013. "How Immigrant Children Affect the Academic Achievement of Native Dutch Children." *Economic Journal* 123(570):F308-F31.
- Parsons, Craig A., and Timothy M. Smeeding. 2006. "What's unique about immigration in Europe?" Pp.1-29 in *Immigration and the Transformation of Europe* edited by Craig A. Parsons and Timothy M. Smeeding. Cambridge: Cambridge University Press.
- Plotnick, Robert D., and Saul D. Hoffman. 1999. "The Effect of Neighborhood Characteristics on Young Adult Outcomes: Alternative Estimates." *Social Science Quarterly* 80:1-18.
- Portes, Alejandro and Rubén G. Rumbaud. 2001. "Legacies: The Story of the Immigrant Second Generation". Berkeley: University of California Press.
- Portes, Alejandro, Patricia Fernandez-Kelly, and William Haller. 2005. "Segmented Assimilation on the Ground: The New Second Generation in Early Adulthood." *Ethnic and Racial Studies* 28: 1000-1040.
- Robertson, Donald, and James Symons. 2003. "Self-selection in the state school system." *Education Economics* 11:259-73.
- Sampson, Robert J., Jeffery D. Morenoff, and Thomas Gannon-Rowley. 2002. "Assessing 'Neighborhood Effects': Social Processes and New Directions in Research." *Annual Review of Sociology* 28:443-78.
- Schwartz, Amy Ellen, and Leanna Stiefel. 2011. "Immigrants and Inequality in Public Schools." Pp. 419-442 in *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, edited by Duncan, Greg J., and Richard J. Murnane. New York/Chicago: Russell Sage Foundation/Spencer Foundation.
- Skolverket (2012). Likvärdig utbildning i svensk grundskola. En kvantitativ analys av likvärdighet över tid. *Rapport 2012: 374*.

- Skolverket (2013). PISA 2012: 15-åringars kunskaper i matematik, läsförståelse och naturvetenskap. *Rapport 2013: 398*.
- Spaiser, Viktoria, Peter Hedström, Shyam Ranganathan, Kim Jansson, Monica K. Nordvik, and David JT Sumpter. 2016. "Identifying Complex Dynamics in Social Systems A New Methodological Approach Applied to Study School Segregation." *Sociological Methods & Research*, 0049124116626174.
- Statistics Sweden (2015). scb.se/statistikdatabasen/befolkning.
- Szulkin, Ryszard, and Jan O. Jonsson. 2007. "Ethnic Segregation and Educational Outcomes in Swedish Lower secondary Schools." *Stockholm University Linnaeus Center for Integration Studies Working Paper 2007:2*.
- Wilson, William J. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.
- Winship, Christopher, and Robert D. Mare. 1984. "Regression Models with Ordinal Variables." *American Sociological Review*, 49: 512–25.
- Winston, Gordon C., and David J. Zimmerman. 2004. "Peer Effects in Higher Education." Pp. 395-421 in *College Choices: The Economics of Where to Go, When to Go and How to Pay for It*, edited by Caroline Hoxby. Chicago: University of Chicago Press for the NBER-C.
- Zhou, Min. 1997. "Growing up American: The Challenge Confronting Immigrant Children and Children of Immigrants." *Annual Review of Sociology* 23:63-95.

Table 1**Descriptive statistics for the main independent and dependent variables**

	Native Swedes of Swedish background					Second generation					First generation immigrants				
	Median	Mean	SD	Min	Max	Median	Mean	SD	Min	Max	Median	Mean	SD	Min	Max
Proportion first generation immigrant pupils	0.03	0.05	0.07	0.00	0.91	0.10	0.14	0.14	0.00	0.91	0.13	0.18	0.17	0.00	0.90
Proportion pupils with academic background	0.37	0.39	0.15	0.00	1.00	0.34	0.37	0.15	0.02	0.97	0.32	0.35	0.14	0.02	0.97
Proportion first generation immigrant pupils from refugee countries	0.01	0.03	0.05	0.00	0.81	0.05	0.09	0.11	0.00	0.81	0.08	0.12	0.14	0.00	0.81
Proportion first generation immigrant pupils who arrived after age 10	0.01	0.02	0.03	0.00	0.70	0.03	0.06	0.07	0.00	0.68	0.05	0.08	0.09	0.00	0.70
Proportion co-ethnics in school	-	-	-	-	-	-	-	-	-	-	0.01	0.03	0.05	0.00	0.42
Proportion non-co-ethnics in school	-	-	-	-	-	-	-	-	-	-	0.10	0.15	0.15	0.00	0.90
Eligible for upper secondary school	-	0.91	0.28	0	1	-	0.84	0.36	0	1.00	-	0.68	0.47	0.00	1.00
Grade point score	210	208	63	0	320	200	195	70	0	320	180	169	80	0	320
N	1362456					106259					105522				

Table 2

OLS regressions on the association between ethnic concentration and grade point scores

	Native Swedes of Swedish background						Second generation						First generation immigrants					
	<i>No controls</i>			<i>With controls</i>			<i>No controls</i>			<i>With controls</i>			<i>No controls</i>			<i>With controls</i>		
	Model 1			Model 2			Model 3			Model 4			Model 5			Model 6		
	b	se	p	b	se	p	b	se	P	b	se	p	b	se	p	b	se	p
Proportion first generation immigrant pupils	-37.47	5.83	0.000	-28.83	5.57	0.000	-39.80	6.75	0.000	-21.45	6.66	0.001	-65.61	8.45	0.000	-41.61	7.37	0.000
Proportion pupils with academic background	68.96	2.73	0.000	21.02	2.57	0.000	51.45	6.00	0.000	26.10	5.63	0.000	35.84	10.35	0.001	11.49	7.94	0.148
Proportion first generation immigrant pupils from refugee countries	-43.34	5.16	0.000	-33.01	4.22	0.000	-46.33	6.81	0.000	-22.44	6.12	0.000	-63.63	5.39	0.000	-40.38	5.12	0.000
Proportion first generation immigrant pupils who arrived after age 10	-81.96	13.78	0.000	-62.04	13.73	0.000	-72.16	14.42	0.000	-43.60	14.50	0.003	-157.76	20.10	0.000	-91.54	17.86	0.000
N	1362456						106259						105522					

Baseline model includes the proportion of children with tertiary educated parents, calendar year, and an indicator of whether the schools is an independent school

Control variables include sex, educational level of parents, living in an intact family, number of siblings, position of the individual in the sibling set, parental employment, household income, age at immigration (for those born abroad)

Table 3**School fixed effects regressions on the association between ethnic concentration and grade point scores**

	Native Swedes of Swedish background			Second generation			First generation immigrants			
	Model 1			Model 2			Model 3			
	b	se	p	b	se	p	b	se	p	
Proportion first generation immigrant pupils	-13.89	3.20	0.000	-2.54	5.31	0.632	-5.73	6.27	0.361	
Proportion pupils with academic background	0.40	1.89	0.834	6.48	4.73	0.171	-2.69	5.56	0.628	
Proportion first generation immigrant pupils from refugee countries	-13.68	4.35	0.002	0.49	6.87	0.943	-5.18	7.74	0.503	
Proportion first generation immigrant pupils who arrived after age 10	-13.98	4.77	0.003	1.45	8.52	0.865	-	13.05	11.03	0.237
N	1362456			106259			105522			
N schools	1339			1330			1327			

All models control for calendar year, independent school, sex, educational level of parents, living in an intact family, number of siblings, position of the individual in the sibling set, parental employment, household income, age at immigration (for those born abroad)

Table 4

Family fixed effects and family and school fixed effects regressions on the association between ethnic concentration and grade point scores

	Native Swedes of Swedish background									First or second generation immigrants								
	<i>Family FE</i>						<i>Family and school FE</i>						<i>Family FE</i>			<i>Family and school FE</i>		
	<i>Stayers</i>			<i>Movers</i>			<i>Movers</i>			<i>Stayers</i>			<i>Movers</i>			<i>Movers</i>		
	Model 1			Model 2			Model 3			Model 4			Model 5			Model 6		
	b	se	p	b	se	p	b	se	p	b	se	p	b	se	p	b	se	p
Proportion first generation immigrant pupils	6.30	2.64	0.017	-10.88	2.72	0.000	-1.46	4.14	0.725	6.70	4.16	0.108	-6.05	2.99	0.043	6.84	5.07	0.177
Proportion pupils with academic background	-3.91	1.30	0.003	-6.48	1.53	0.000	2.10	2.46	0.392	-10.71	4.21	0.011	-5.17	3.41	0.129	-4.44	5.77	0.442
Proportion first generation immigrant pupils from refugee countries	6.03	3.52	0.087	-2.23	3.38	0.510	3.16	5.42	0.560	11.03	5.00	0.027	-6.48	3.58	0.071	5.31	5.99	0.375
Proportion first generation immigrant pupils who arrived after age 10	6.29	3.93	0.110	-36.03	5.64	0.000	-7.37	6.33	0.244	14.02	6.19	0.024	-15.77	5.43	0.004	6.08	7.27	0.403
N	1138283			224173			224173			158651			53130			53130		
N families	741345			93836			93836			106905			20016			20016		
N schools							1339									1296		

All models control for calendar year, independent school, proportion of children with tertiary educated parents, sex, educational level of parents, living in an intact family, number of siblings, position of the individual in the sibling set, parental employment, household income, age at immigration (for those born abroad)

Table 5

OLS regressions on the association between ethnic concentration and eligibility

	Native Swedes of Swedish background						Second generation						First generation immigrants					
	<i>No controls</i>			<i>With controls</i>			<i>No controls</i>			<i>With controls</i>			<i>No controls</i>			<i>With controls</i>		
	Model 1			Model 2			Model 3			Model 4			Model 5			Model 6		
	b	se	p	b	se	p	b	se	p	b	se	p	b	se	p	b	se	p
Proportion first generation immigrant pupils	-0.24	0.02	0.000	-0.21	0.02	0.000	-0.33	0.03	0.000	-0.25	0.03	0.000	-0.48	0.03	0.000	-0.34	0.03	0.000
Proportion pupils with academic background	0.15	0.01	0.000	0.04	0.01	0.000	0.17	0.02	0.000	0.09	0.02	0.000	0.14	0.04	0.000	0.04	0.03	0.181
Proportion first generation immigrant pupils from refugee countries	-0.28	0.02	0.000	-0.24	0.02	0.000	-0.39	0.04	0.000	-0.29	0.04	0.000	-0.48	0.03	0.000	-0.36	0.03	0.000
Proportion first generation immigrant pupils who arrived after age 10	-0.48	0.05	0.000	-0.41	0.05	0.000	-0.55	0.07	0.000	-0.43	0.07	0.000	-1.05	0.08	0.000	-0.65	0.06	0.000
N	1362456						106259						105522					

Baseline model includes calendar year, proportion of children with tertiary educated parents and independent school

Control variables include sex, educational level of parents, living in an intact family, number of siblings, position of the individual in the sibling set, parental employment, household income, age at immigration (for those born abroad)

Table 6**School fixed effects regressions on the association between ethnic concentration and eligibility**

	Native Swedes of Swedish background			Second generation			First generation immigrants		
	Model 1			Model 2			Model 3		
	b	se	p	b	se	p	b	se	p
Proportion first generation immigrant pupils	-0.10	0.01	0.000	-0.05	0.03	0.104	-0.11	0.04	0.002
Proportion pupils with academic background	0.02	0.01	0.004	0.05	0.02	0.035	0.02	0.03	0.455
Proportion first generation immigrant pupils from refugee countries	-0.10	0.02	0.000	-0.05	0.04	0.181	-0.09	0.04	0.050
Proportion first generation immigrant pupils who arrived after age 10	-0.11	0.02	0.000	-0.04	0.05	0.450	-0.13	0.06	0.024
N	1362456			106259			105522		
N schools	1339			1330			1327		

All models control for calendar year, independent school, proportion of children with tertiary educated parents, sex, educational level of parents, living in an intact family, number of siblings, position of the individual in the sibling set, parental employment, household income, age at immigration (for foreign born)

Table 7

Family fixed effects and family and school fixed effects regressions on the association between ethnic concentration and eligibility

	Native Swedes of Swedish background									First or second generation immigrants								
	<i>Family FE</i>						<i>Family and school FE</i>						<i>Family FE</i>			<i>Family and school FE</i>		
	<i>Stayers</i>			<i>Movers</i>			<i>Movers</i>			<i>Stayers</i>			<i>Movers</i>			<i>Movers</i>		
	Model 1			Model 2			Model 3			Model 4			Model 5			Model 6		
	b	se	p	b	se	p	b	se	p	b	se	p	b	se	p	b	se	p
Proportion first generation immigrant pupils	-0.04	0.02	0.017	-0.13	0.01	0.000	-0.05	0.02	0.039	-0.01	0.03	0.792	-0.11	0.02	0.000	-0.08	0.03	0.017
Proportion pupils with academic background	0.01	0.01	0.132	0.01	0.01	0.523	0.04	0.01	0.010	0.02	0.03	0.566	0.03	0.02	0.229	0.02	0.04	0.689
Proportion first generation immigrant pupils from refugee countries	-0.03	0.02	0.111	-0.11	0.02	0.000	-0.01	0.03	0.662	0.03	0.03	0.464	-0.12	0.02	0.000	-0.09	0.04	0.018
Proportion first generation immigrant pupils who arrived after age 10	-0.05	0.02	0.027	-0.29	0.03	0.000	-0.12	0.04	0.001	0.01	0.04	0.783	-0.19	0.03	0.000	-0.04	0.05	0.445
N	1138283			224173			224173			158651			53130			53130		
N families	741345			93836			93836			106905			20016			20016		
N schools							1339									1296		

All models control for calendar year, independent school, proportion of children with tertiary educated parents, sex, educational level of parents, living in an intact family, number of siblings, position of the individual in the sibling set, parental employment, household income, age at immigration (for those born abroad)

Table 8

OLS regressions and fixed effects regressions on the association between co-ethnics and school outcomes

	<i>OLS regressions</i>			<i>School FE</i>			<i>Family FE</i>						<i>Two-way FE</i>		
							<i>Stayers</i>			<i>Movers</i>			<i>Movers</i>		
	Model 1			Model 2			Model 3			Model 4			Model 5		
	b	se	p	b	se	p	b	se	p	b	se	p	b	se	p
Grade point score															
Proportion non-co-ethnics in school	-37.05	1.63	0.000	-7.13	3.60	0.048	-6.99	7.31	0.339	-20.52	4.92	0.000	-2.55	8.94	0.776
Proportion co-ethnics in school	-59.41	5.07	0.000	0.73	6.08	0.904	55.08	13.95	0.000	33.28	13.83	0.016	54.79	16.24	0.001
Proportion pupils with academic background	7.81	1.77	0.000	-4.12	3.91	0.292	-6.19	7.09	0.383	-6.21	5.48	0.257	-3.64	9.65	0.706
Eligibility															
Proportion non-co-ethnics in school	-0.30	0.01	0.00	-0.12	0.02	0.000	-0.10	0.05	0.044	-0.17	0.03	0.000	-0.08	0.06	0.198
Proportion co-ethnics in school	-0.43	0.03	0.00	-0.07	0.04	0.075	0.44	0.10	0.000	0.00	0.09	0.983	0.15	0.11	0.172
Proportion pupils with academic background	0.04	0.01	0.00	0.02	0.02	0.356	0.04	0.05	0.468	0.00	0.04	0.955	-0.05	0.07	0.446
N	105484			105484			78525			26959			26959		
N schools				1327									1218		
N families							57717			12496			12496		

All models control for calendar year, independent school, proportion of children with tertiary educated parents, sex, educational level of parents, living in an intact family, number of siblings, position of the individual in the sibling set, parental employment, household income, age at immigration (for those born abroad), and country/region of birth

END NOTES

¹ We have access to measures of mean earnings and the proportion of employed parents at the school level; however, these variables are too highly correlated with the proportion of immigrants to be included in our models.

² We have also tested alternative measures of ethnic concentration, such as the proportion of immigrants who arrived in Sweden at age 13 or later and the proportion of immigrants from refugee countries who arrived in Sweden after the age of 6 (school start). In all of these analyses, the associations between the indicators of school segregation and grades are stronger. However, the narrower the definition of segregation that is used, the smaller is the within-school and within-family variation of the variable.

³ We also note (Models 1, 2, 4 and 5) the negative effects of socio-economic school segregation (proportion of students from an academic background), yet the coefficients are small and not significant in the two-way fixed effect models. The largest coefficient, -10.71 for immigrants, implies that if an immigrant changed (this estimate is valid for school stayers) from a school cohort with zero students of academic background to a school cohort where 100 percent of students were of academic background, his/her expected grade would decrease by about 11 points (on the grade score scale that ranges from 0 to 320). This is nevertheless a counter-intuitive finding which might be due to teachers being less willing to use high grades in school classes with large numbers of children from an academic background.

Appendix A

Countries and regions in data

Country	N	Country	N
Sweden	1 445 957	USA and Canada	1 466
Iraq	18 085	Ethiopia	1 418
Yugoslavia	11 832	China, Mongolia, and North Korea	1 416
Bosnia and Herzegovina	11 003	Romania	1 408
Iran	7 314	British Islands and Oceania	1 385
India region	4 799	Central America, the Caribbean, and Mexico	1 310
Poland	4 152	Philippines and Polynesia	1 208
Somalia	4 102	West Africa	1 107
Thailand	3 601	Central Asia and Caucasus	1 066
Colombia	3 382	Serbia and Montenegro	951
Chile, Argentina, and Uruguay	3 162	Southeast Asia (other)	919
Lebanon	3 036	Benelux	835
German Europe	2 927	Central Africa	807
High income Asia	2 849	Eritrea	726
Russia / Soviet	2 759	Italian / French Europe	718
Turkey	2 683	Lithuania and Latvia	688
Afghanistan	2 613	Hungary	539
South America (other)	2 571	Estonia	440
Norway	2 396	Tunisia, Libya, Egypt	429
Denmark and Iceland	2 104	Iberia and Pyrenees	424
Syria	1 968	Pakistan	396
Vietnam	1 961	Morocco and Algeria	349
Finland	1 882	Greece and Cyprus	347
Middle East (other)	1 873	Missing / unknown	49
the Victoria Region and Southern Africa	1 660		
Eastern Europe (other)	1 625		
Yugoslavia (other) and Albania	1 540		

Appendix B

Within-family variation in the main independent and dependent variables

Table B1 describes the within-family variation in our main independent variables, by subtracting the lowest value in the sibling set from the highest value in the sibling set. In the interest of saving space, we do not distinguish between different immigrant backgrounds in Table B1, with the exception of the last column ('Proportion co-ethnics') which only includes first generation immigrants.

The within-family variation is indeed substantial. 93 percent of all sibling sets have some variation in the proportion of immigrants at school, and the mean difference in these sibling sets is 4.6 percentage points. If siblings attend the same school (those whom we term 'stayers'), the mean difference between the siblings is 3 percentage points. This difference is considerably larger, however, (8 percentage points) if the siblings attend different schools (those whom we term 'movers'). Differences between siblings are, naturally, consistently larger among movers than among stayers. The mean sibling difference in the proportion of pupils with an academic background is around 6 percentage points for movers and 13 percentage points for stayers. The mean differences between siblings in the proportion of recently arrived immigrants and in the proportion of immigrants from typical refugee countries follow the same logic as that described for the proportion of immigrants; however, the mean sibling difference is consistently smaller. For first generation immigrants, we have also calculated the mean sibling difference in the proportion of co-ethnics in school. Only 57 percent of the sibling sets include some variation on this variable. This is because a large proportion of the immigrants have no co-ethnics at school (40 percent). Again, we find larger sibling differences among movers, and smaller differences among stayers.

Table B1: Within-family differences in school characteristics

	School characteristics														
	Proportion first generation immigrant pupils			Proportion pupils with academic background			Proportion first generation immigrant pupils who arrived after age 10			Proportion first generation immigrant pupils from refugee countries			Proportion co-ethnics (only first generation immigrants)		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Percentage sibling sets with within-family difference	93.3%			97.9%			80.0%			80.8%			56.7%		
Mean within-family difference	0.043	0.000	0.899	0.078	0.000	0.762	0.024	0.000	0.696	0.030	0.000	0.813	0.020	0.037	0.422
Mean within-family difference, if >0	0.046	0.000	0.899	0.080	0.000	0.762	0.030	0.000	0.696	0.038	0.000	0.813	0.035	0.000	0.422
Mean within-family difference, if >0 and sibling set attend same school (stayers)	0.033	0.000	0.636	0.062	0.000	0.458	0.024	0.000	0.473	0.028	0.000	0.562	0.029	0.000	0.422
Mean within-family difference, if >0 and sibling set attend different schools (movers)	0.084	0.000	0.899	0.132	0.000	0.762	0.046	0.000	0.696	0.065	0.000	0.813	0.046	0.000	0.386

Table B2: Within-family differences in siblings' school outcomes

		<u>Eligibility</u>	<u>GPA</u>
Percentage sibling sets with within-family difference	All	14.0%	96.4%
	If sibling set attend same school	12.2%	96.1%
	If sibling set attend different schools	19.6%	97.3%
Mean within-family difference	All	-	53.1
	If >0	-	55.1
	If >0 and sibling set attend same school	-	52.0
	If >0 and sibling set attend different schools	-	65.0

Table B2 includes descriptive statistics on how the two variables measuring school performance differ within sibling sets. 14 percent of all sibling sets have at least one sibling who is eligible while another sibling is not. This is particularly common if siblings attend different schools, indicating differences in quality between schools. Almost all sibling sets have some variation in grades, and again, this is particularly common if at least one of the siblings attends a different school. The mean within-family difference in grades is 13 points higher ($65-52=13$) if siblings attend different schools as compared to when the siblings attend the same school. This is equivalent to the difference between a pass and a pass with distinction in almost three out of sixteen subjects.

Appendix C

Descriptive statistics for control variables

	Native Swedes of Swedish background				Second generation				First generation immigrants			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Girl	0.49	0.50	0	1	0.49	0.50	0	1	0.48	0.50	0	1
Sibling order												
1st child	0.49	0.50	0	1	0.41	0.49	0	1	0.48	0.50	0	1
2nd child	0.35	0.48	0	1	0.32	0.47	0	1	0.29	0.45	0	1
3rd child	0.12	0.33	0	1	0.16	0.36	0	1	0.13	0.33	0	1
4th child or later	0.04	0.19	0	1	0.11	0.31	0	1	0.10	0.30	0	1
Number of siblings												
1 sibling	0.14	0.35	0	1	0.12	0.33	0	1	0.15	0.36	0	1
2 siblings	0.47	0.50	0	1	0.34	0.47	0	1	0.29	0.46	0	1
3 siblings	0.28	0.45	0	1	0.27	0.44	0	1	0.22	0.41	0	1
4 siblings	0.08	0.27	0	1	0.15	0.36	0	1	0.14	0.35	0	1
5 or more siblings	0.03	0.17	0	1	0.12	0.33	0	1	0.19	0.39	0	1
Lives in intact family	0.63	0.48	0	1	0.62	0.49	0	1	0.57	0.49	0	1
Parents' education												
Comprehensive school	0.04	0.21	0	1	0.18	0.39	0	1	0.24	0.43	0	1
Upper secondary school	0.48	0.50	0	1	0.48	0.50	0	1	0.35	0.48	0	1
Tertiary education <2 years	0.07	0.26	0	1	0.04	0.20	0	1	0.03	0.17	0	1
Tertiary education 2 years or longer	0.40	0.49	0	1	0.28	0.45	0	1	0.34	0.47	0	1
Missing	0.00	0.03	0	1	0.01	0.07	0	1	0.05	0.21	0	1

Appendix C, continued

Mother is earner	0.82	0.39	0	1	0.59	0.49	0	1	0.40	0.49	0	1
Father is earner	0.80	0.40	0	1	0.52	0.50	0	1	0.34	0.48	0	1
Household income												
Q1	0.14	0.35	0	1	0.42	0.49	0	1	0.65	0.48	0	1
Q2	0.20	0.40	0	1	0.26	0.44	0	1	0.19	0.39	0	1
Q3	0.21	0.41	0	1	0.15	0.36	0	1	0.08	0.28	0	1
Q4	0.22	0.41	0	1	0.11	0.31	0	1	0.04	0.20	0	1
Q5	0.23	0.42	0	1	0.06	0.24	0	1	0.03	0.16	0	1
Missing	0.00	0.03	0	1	0.00	0.05	0	1	0.01	0.09	0	1
Mover												
All siblings attend same school	0.84	0.37	0	1	0.75	0.43	0	1	0.74	0.44	0	1
Siblings attend different schools	0.16	0.37	0	1	0.25	0.43	0	1	0.26	0.44	0	1
Eligible for upper secondary school	0.91	0.28	0	1	0.84	0.36	0	1	0.68	0.47	0	1
Grade point score	208	63	0	320	195	70	0	320	169	80	0	320
Age at immigration												
Before age 7	-	-	-	-	-	-	-	-	0.38	0.48	0	1
Age 7-9	-	-	-	-	-	-	-	-	0.21	0.41	0	1
Age 10-12	-	-	-	-	-	-	-	-	0.21	0.41	0	1
Age 13+	-	-	-	-	-	-	-	-	0.20	0.40	0	1
Proportion co-ethnics in school	-	-	-	-	-	-	-	-	0.03	0.05	0	0.42
Proportion non co-ethnics in school	-	-	-	-	-	-	-	-	0.15	0.15	0	0.90
N	1362456 (86.5%)				106259 (6.7%)				105522 (6.7%)			