The Aspirations-Attainment Paradox of Immigrant Children: A Social Networks Approach*

Julia Boguslaw†

Draft version: August 20, 2017

Abstract

Using two independent and nationally representative samples of Swedish children, I compare the university aspirations and expectations between children of immigrants and children of natives. In line with existing findings, I find that children with foreign-born parents have significantly higher aspirations and expectations than their native-majority peers with and without conditioning on school performance, academic potential and friendship networks. I do not I find evidence of an aspirations-expectations gap among immigrant children: their aspirations and expectations are not less aligned than their native-majority peers. The results reveal significant gender differences. Native-majority girls are for instance more likely to express an aspirations-expectations gap among students who are showing signs of academic potential. Moreover, having only female friends makes one less likely to belong to the aforementioned category.

Keywords: aspirations, expectations, immigrant children, social networks

1 Introduction

Although higher education is free in Sweden children with an immigrant background lag behind children of native-born in educational performance; they have lower grades and are more likely to have incomplete grades or dropout. Empirical findings in the field of educational inequality suggest a number of plausible mechanisms of this result and one important factor is individual aspirations and expectations of higher education (see for example Guyon & Huillery (2016) and Jonsson & Rudolphi (2011) for an literature review). Importantly, immigrant children may face specific challenges in turning aspirations into achievement, for instance sanctions by peers for succeeding in school, structural barriers on the labor market and occupational discrimination.

In spite of these obstacles (or perhaps in part because of them) children of immigrants tend to have higher aspirations than their native majority peers. A significant body of research (Kao & Tienda 1995, Heath & Brinbaum 2007, Jonsson & Rudolphi 2011) has found that ethnic minority or immigrant children generally have more positive attitudes towards education than their native-born counterparts. [1] Black students in the US for instance have been shown to perform worse in school than do white students but still have more positive attitudes toward school than their white peers (Akerlof & Kranton 2002). This phenomenon has been called "the aspiration-attainment paradox of immigrants" (see for example Mickelson (1990) for a discussion).

In this paper, I explore "the aspiration-attainment paradox" by comparing the discrepancies between aspirations for a university degree and the expectations of getting one between children with immigrant parents and children of native-born. Drawing on the theory of Ray (2006), Genicot & Ray (2017) and Dalton et al. (2016) on the "aspirations window" and "aspirations failure", I use the mismatch between educational aspirations and expectations, what I will refer to as "the aspirations-expectations gap" (or more simply "the gap") as a potential mechanism behind the aspirations-attainment gap among immigrant students.[2]

A closely related topic is the concept of "lost talent", a term coined by Hanson (1994) in a seminal paper on the mismatch of aspirations and expectations among American youths showing early signs of academic potential.[3] Inspired by Hanson

---

[1]Jonsson & Rudolphi (2011) do not study attitudes *per se*, but interpret actual transitions to academic tracks as positive attitudes.

[2]See also Genicot & Ray (2009).

[3]Hanson (1994) finds that not race but social class is the strongest predictor of lost talent. One of the critics of this literature, Mickelson (1990), argues that this is a consequence of researchers inability to distinguish between abstract and concrete attitudes. Abstract attitudes are general beliefs about education, for instance the generally held belief that education is important for

([1994](#)), I study the observable characteristics of the pool of students in a representative sample of eight graders in Sweden who are showing signs of early academic potential and who are expressing an aspirations-expectations gap. Finally, I test whether children of immigrants are more likely to belong to this category than are the children of native-born.

The empirical analysis is based on two datasets: the Swedish Level of Living Survey 2010 – Immigrants and their children (LNU-UFB) and Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU, Kalter et al. (2013)). Both datasets are comprised of nationally representative samples of Swedish students and include an oversampling of children with an immigrant background. The stratified samples of LNU-UFB 2010 and CILS4EU allow detailed analyses of the social integration of immigrant children specifically, a group of great interest given the increased importance of immigration in Western countries. The survey data in LNU-UFB 2010 is combined with rich register data on the background characteristics of parents' such as region of birth.[4]

The analysis is structured as follows. First, I investigate students' aspirations and expectations with respect to their parents' immigration status. The outcome variables *educational aspirations* and *educational expectations* are drawn from the following two questions found in the LNU survey: "Would you like to continue going to school after the upper secondary level, that is, attend a university or university college?" and "Do you think you actually will continue going to school after the upper secondary level?" and the corresponding questions in CILS4EU: "What is the highest level of education you wish to get? (Don't know, No degree, Compulsory school, Upper secondary school, College/university)" and "What is the highest level of education think you will actually get? (Don't know, No degree, Compulsory school, Upper secondary school, College/university)". The variable *aspirations-expectations gap* is constructed using the variables *educational aspirations* and *educational expectation* and is defined as having aspirations for an university degree but not expecting of getting one.

I try to account for potential differences using information at the family, friendship, and classroom level. By using the multilevel-structured dataset CILS4EU, I try to gauge the influence of the gender, immigrant status and socioeconomic background composition within an individual's network of best friends on the three outcome variables educational aspirations, expectations and the aspirations-

---

socioeconomic mobility, while concrete attitudes are formed through actual experiences. According to Mickelson (1990), concrete rather than abstract attitudes determine school performance.

[4]Apart from Rudolphi (2014), the LNU child survey has previously been used in a study by Jonsson & Östberg (2009) and the governmental report of Mood & Jonsson (2013). The study Olsson (2009) also uses Child-LNU data to look at the role of social relations for disadvantaged adolescents.

expectations gap. As a final step, I explore the observable characteristics of students scoring higher than the mean or median of the sample on a standardized cognitive test in grade 8 and who are expressing an aspirations-expectations gap (henceforth referred to as "lost talent").

This study contributes to the emerging economic literature on aspirations (Goux et al. forthcoming, Carlana et al. 2015, Guyon & Huillery 2016) and disparities in life courses of individuals. To the best of my knowledge, this is the first study that explores the predictors of "lost talent" among children with foreign-born parents in a European context.[5] A handful of studies have tried to explain differences in aspirations or expectations among different socio-economic or ethnic groups (Salikutluk 2016, Hanson 1994, Rudolphi 2014, Heath & Brinbaum 2007).[6] This study contributes to the existing literature by studying the gap between aspirations and expectations from a social network approach using self-reported friendship links. Unlike the studies of Rudolphi (2014) and Guyon & Huillery (2016) focusing on low-SES children, this paper utilizes a dataset that includes an oversampling of foreign-born children which facilitates a broader analysis of disadvantage. Of particular interest is the interaction between immigrant background and gender.

It also contributes to the literature on dropout decision (Goux et al. forthcoming, Mora & Oreopoulos 2011) by exploring the aspirations-expectations gap among low-performing students at risk of dropout after compulsory school. A closely related paper is that of Guyon & Huillery (2016) which estimates the influence of aspirations on school outcomes and empirically tests the aspirations failure model of Ray (2006). They find a social gradient in aspirations failure (operationalized using survey questions about high school track choice, awareness of existing tracks and self-perceived academic potential), validating previous findings (e.g. Hanson (1994)). In contrast to Guyon & Huillery (2016) and Salikutluk (2016) which are based on either relatively small or unrepresentative survey samples, this study uses two independent and nationally representative samples of Swedish youths combined with detailed friendship network data.

Aspirations are determined through social interactions. The most relevant social networks studies for the purposes of this paper are Guyon & Huillery (2016), Burgess & Umaña-Aponte (2011) and Mora & Oreopoulos (2011) of which the latter two estimate the impact of the social network on individual educational aspirations (for example whether or not to drop out from school).[7] So far, research on social

---

[5]The only paper on lost talent that I am aware of is the study of Hanson (1994).

[6]Salikutluk (2016) tests the explanatory power of different explanations for ethnic disparities in aspirations for upper-secondary education using survey data on immigrant youth in Germany. Based on a survey of 1,052 cases the results show that the theories of immigrant optimism, blocked opportunities and social capital contribute to explain the ethnic gap in aspirations.

[7]Burgess & Umaña-Aponte (2011) find significant effects of the socio-economic background of

networks and educational decisions have been confined to general peer effects and less attention has been paid to the interaction of immigrant status and gender. This paper fills the existing gap by studying the influence of friends' characteristics on educational plans.

Based on two independent and nationally representative samples, I find that children with an immigrant background generally have higher aspirations and expectations than the majority group. Overall, I do not find evidence of a significant aspirations-expectations gap among immigrant children: their aspirations and expectations are not more likely to be unaligned than their native-majority peers'. I find significant gender differences: girls are more likely to express an aspirations-expectations gap and be labeled as lost talent. I find that gender and language proficiency are the strongest predictors of lost talent. Moreover, having only female friends makes one less likely to belong to the aforementioned category.

The rest of this paper is structured as follows. In section 2, I give an overview of previous research and the relevant terms and concepts. Data and definitions are presented in section 3. In section 4-6, I show the results and discuss their robustness. Finally, in section 7, I discuss the policy implications of my findings and give some concluding comments.

2 Theoretical framework

2.1 Aspirations, expectations and human capital formation

Educational aspirations and expectations are important since they can help to explain and predict individual differences in educational choices.[8] According to human capital theory (Becker 1964, Mincer 1974, Schultz 1960, Björklund et al. 2014), individuals make educational choices based on the calculations of the future wage returns from investments in human capital. Expectations reflect an individual's plans of investment in human capital. Given that individuals base their decisions on calculations of future wage returns, those with the highest rates of return to education should also have the highest educational expectations (see for example Morgan (1998)).

A highly relevant study for the purpose of this paper is the study of Hanson (1994) which looks at the mismatch between educational aspirations and expecta-

---

friends for own educational aspirations and expectations. See also Roth & Salikutluk (2012) on aspirations and social networks of mothers. The results of Mora & Oreopoulos (2011) suggest that the influence of non-reciprocating friends' dropout plans for an individual's decision to dropout is small and insignificant.

[8]See for example Feliciano & Rumbaut (2005), Jacob & Wilder (2010), Morgan (2005), and Portes & Rumbaut (2001).

tions to measure the amount of "lost talent" among American youths measured in senior years of high-school. The sample is restricted to individuals with high educational expectations and above average scores on standardized Mathematics and reading tests. Hanson (1994) finds that 16 percent of the sample of youths who aspire to a college degree do not expect to attain one. Furthermore, she finds no significant effect of race in logistic regression models predicting lost talent. The strongest predictor of lost talent is social class.

The educational aspirations and expectations of Swedish children have previously been studied by Rudolphi (2014) who looks at the consistency between wanting to continue going to school after upper secondary level (aspirations) and thinking one will actually attend university studies (expectations) in a sample of compulsory school students in Sweden (LNU, n=620). The results reveal that students generally show consistency between high aspirations and high expectations. Also, girls tend to show more consistency than do boys. The study does not look specifically at immigrant children.[9]

2.2 Social determinants of aspirations and expectations

A much less explored topic within the literature on aspirations and expectations is the role of peer effects for students' educational plans. Previous studies on social networks and peer effects have shown that reference groups and friends are important determinants of scholastic achievement (e.g. Calvó-Armengol et al. (2009)). The theoretical paper of Ray (2006) incorporates the social dimension of aspirations in his so-called "aspirations-based view of individual behavior" arguing that individual preferences and behavior depend *both* on historical experience *and* the experiences of individuals that are similar or close to him or her (socially and spatially). Furthermore, aspirations are determined through social interactions and are transmitted to children from parents, friends, classroom peers and adults in the community in which they live (Appadurai 2004). The empirical papers of Burgess & Umaña-Aponte (2011) and Mora & Oreopoulos (2011) mentioned previously estimate the impact of friendship network on individual educational aspirations and drop out intentions using cross-sectional network data.

The positive effects of parental tutoring on dropout behavior has for example

---

[9]The studies of Halleröd (2011) and Alm (2011) examine the relationship between different outlooks in adolescence and later outcomes in life using Swedish data. The findings of Halleröd (2011) indicate that there is a relationship between children's expectations in the ages 12-13 and outcomes later in adult life: children who had a more pessimistic view on their future (approach to the future, future outlooks, expectations, beliefs about the future) were more at risk of economic hardship and weak labor market attachment. The results of Alm (2011) show that indifference in adolescence is positively correlated with economic hardship and low educational attainment later in life.

been demonstrated by Goux et al. (forthcoming) who use a randomized control trial on parents' aspirations. The treated sample, the parents of low-performing students with unrealistic aspirations, was provided extra counseling with respect to their children's transition to high school. The intervention was not planned to improve these students' school results but to attune their aspirations to their school performance so as to reduce the risk of drop out in high school due to unrealistic educational plans. A similar RCT has previously been conducted by (Avvisati et al. 2014) where the treatment consisted of informational meetings between parents and the school head. The parents were given advice on how to assist their children with their school work and the findings indicate that parental involvement had an significant impact on student behavior.

As schools with a high proportion of foreign-born students tend to be located in areas of concentrated economic disadvantage with neighborhoods composed of adults with low educational and labor market aspirations, studying the role of social networks is key for understanding the determinants of ethnic disparities in educational decisions.[10] Another relevant study is Carlana et al. (2015) which presents positive causal estimates of the influence of motivational meetings on the choice of high school program among high performing immigrant boys in Italy. In 2011, the Italian Ministry of Education initiated an educational program "to induce students to undertake educational decision congruous to their potentialities". The intervention was designed as a randomized control trial sampling schools and students in northern Italy, specifically targeting high performing immigrant students. The purpose was to adjust students aspirations before choice of high school program and to make high performing immigrant students more likely to consider an academic track in upper-secondary school.

2.3 The aspirations gap and poverty traps

The concept of aspirations can also be found in the more recent inequality literature, for example in the theoretical papers of Ray (2006), Genicot & Ray (2017) and Dalton et al. (2016), which explore the link between aspirations and poverty. Dalton et al. (2016) develop a model according to which external constraints make poor individuals more susceptible to behavioral poverty traps or so-called "aspirations failure". Within this framework, individuals can face both internal (identity) or external constraints (credit/budget constraints) and Dalton et al. (2016) refers to this as a behavioral bias in setting aspirations.[11] For example, the rate of return to higher eduction of immigrant children may be lower than that of their native-born

---

[10]See for example "the memberships theory of inequality" of Durlauf (2006).

[11]See also the sociological literature on internal/external locus of control.

peers. Immigrant children may be well-aware of the benefits of higher eduction but the individual returns are perceived as low due to for example labor market discrimination and lack of role models in their community or neighborhood.

Genicot & Ray (2017) introduce the concept of an "aspirations window" to describe the group of individuals who serve as a reference point, the social frame that shapes an individual's aspirations. The authors argue that aspirations and the social environment in which they are determined should be incorporated into standard economic theory. In their paper they relate aspirations to the income distribution, investment and growth on a macro level arguing that optimal aspirations are those "that lie at a moderate distance from the individual's current economic standard, large enough to incentivize but not so large as to induce frustration". A more recent contribution to this literature is the empirical study by Guyon & Huillery (2016) which tests their model against real data.

On the same topic, Ray (2006) presents a model where individual behavior is determined by the gap between one's current economic standard and the aspired one. According to this idea individuals who have small aspirations gaps are less motivated to raise their standard of living. The same applies to individuals who have wide gaps: they have low incentives to change their standard of living because the necessary investment is too large. Ray (2006) cautions that if the aspirations window is too wide it can result in what the he describes as "the curse of frustrated aspirations", a state in which the incentives to make an investment effort are low and the cost of narrowing the gap is high relative to the benefits. This implies that investment efforts are smallest among individuals with relatively high or low aspiration gaps.[12]

3 Data and definitions

This paper makes use of two datasets both oversampling children with an immigrant background: the Swedish Level of Living Survey 2010 – Immigrants and their children (LNU-UFB) and Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU, Kalter et al. (2013)). Both datasets comprise a nationally representative samples of Swedish children and are described in detail below.

---

[12]The incidence of aspirations failure depends on the degree of polarization in a society. Ray (2006) presents two types of societies where the first is connected and the other is polarized. The aspirations window is wider in societies where perceived mobility is high. Regardless of whether the poor include the rich in their aspirations window, increased polarization will lead to aspiration failure. Thus, aspirations can both incentivize and create frustration: if too high, aspirations can discourage effort investments in for example education. Furthermore, frustration can give rise to what sociologist and anthropologists call an "oppositional culture" (see for example Fordham & Ogbu (1986), which is expressed through weak scholastic achievements and misbehavior (Parrillo 2008).

3.1 The Swedish Level of Living Survey (LNU)

The Swedish Level of Living Survey (LNU) is a panel survey of the regular level of living of the Swedish population and has been conducted around every tenth year since 1968. In 2010 the sample was extended with a subsample the foreign-born population in Sweden. Thus, LNU-UFB is a survey of foreign-born individuals and their children in Sweden. It comprises of foreign-born individuals with a permanent residence permit in Sweden and their children. The respondents (parents) have been living in Sweden for at least five years. The respondent's children were either born in Sweden or abroad and are 10 to 18 years old and living at home.

The LNU-UFB survey has the same design as the Swedish Level of Living Survey (LNU) of 2010. Both surveys were carried out by Statistics Sweden during the years 2010 to 2012 and included interviews with a total of 1,357 respondents in the ages 10 to 18 living at home. The surveys were run simultaneously and the content of the questionnaires were identical which means that LNU-UFB has a reference group consisting of the children of a representative sample of the Swedish population. LNU-UFB contains responses from a total of 437 respondents and the LNU 920. The children filled in a questionnaire by listening to recorded questions with a tape recorder. The child survey is comprised of questions in the following areas: children's material and financial resources, health, household work, neighborhood characteristics and education.

The LNU child sample is restricted to children in the ages 13-18 (grade 7 and over) living at home. Descriptive statistics for the analysis sample is provided in table 1. The sample contains 874 child respondents, with either one, two or no native-born parent. Each child is paired with a parent included in either LNU or LNU-UFB. In order to match the dataset from the LNU-UFB to the parents, all the identical twins have been removed (four pairs) since it is impossible to distinguish between same sex identical twins. LNU and LNU-UFB contain data on respondents' income, employment biography, education history, and health. LNU-UFB is based on a stratified sample of foreign-born individuals hence the survey results are weighted by the parent's region of birth and age (a representative sample of foreign born in the age span 18-75, who have lived in Sweden for at least 5 years). LNU-UFB provides information on the the region of origin of the parent and his or her partner.

I match the LNU-UFB sample with register data from LISA (Longitudinal integration database for health insurance and labor market studies) to obtain reliable income measures and additional information on the parents background, for example region of birth of the partner.[13] LISA (formerly known as LOUISE) was constructed

---

[13]Information on the partner is missing in LNU 2010, which is why I resort to register data on

by Statistics Sweden, the Social Insurance Agency and the Swedish Agency for Innovative Systems and consists of annual registers since 1990. It includes all individuals 16 years of age and older that were registered in Sweden as of December 31 each year.

3.2 Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU)

The second dataset used in this study, CILS4EU is a new, longitudinal cohort survey conducted in four countries: the UK (only England), Germany, the Netherlands, and Sweden. CILS4EU is a multileveled survey containing rich information on the family, teacher, school and classroom. It includes five sub-questionnaires directed at students, parents and teachers: "Youth main", "Youth classmates","Parents", "Teachers" and "Youth friends". In this study, I make use of "Youth main", "Youth classmates" and "Parents". The sample employs a two-stage stratified cluster design, interviewing students in sampled school classes. While the sample is designed to be nationally representative, it features a deliberate overweighting of schools with many children of immigrant background.

The first wave of CILS4EU was conducted in the school year 20102011. Participating students were in the eighth grade (aged 14-15). The number of respondents in the main questionnaire in Sweden in the school year 2010-2011 was 5,025. Amongst these a total of 4,804 students took both the language and cognitive test. The network analysis sample is constructed in the following way. In order to calculate the average characteristics in an individual's friendship network I remove all those that have missing values on any of the variables female, foreign, parents' education, educational aspirations and expectations. Next, I add the friendship network data and thereafter I am left with a total of 4,364 observations. Finally, I match these data with the parent questionnaire which contains information on parents' aspirations and expectations for their children's educational attainment.

The second wave of CILS4EU was administered in 2011-2012 when the respondents attended ninth (aged 15-16) grade. The survey questions regarding aspirations and expectations are identical in the two waves. Information on cognitive and language test results are found in the first wave of the survey while self-reported grades in the core subjects Maths, Swedish and English are included in the second wave. Due to attrition and item non-response the sample size shrinks to 3,631 individuals from wave 1 to wave 2.[14] The relevant variables for this study are presented in table

---

partner's country of origin.

[14]The number of unmatched individuals is 1,633 (total from wave 1 and wave 2). In total 900 individuals are lost due to attrition.

3.3 Friendship network

I use the Swedish sociometric classroom data (n=4,794) which was collected in the first wave of CILS4EU. The friendship network questionnaire was in many of the cases administered during a class with the homeroom teacher. Friendship is defined on the basis of the question "Who are your best friends in this class?" to which the student could nominate a maximum of five individuals. A link between two students exists if an individual nominated another as a "best" friend.

Students who were absent on the day of the network questionnaire or who refused to participate were excluded from the school class roaster and the set of potential friend nominees. Individuals with no friends have been dropped from the friendship network analysis (in total 266 individuals from the analysis sample of non-missing cases defined above).[15] Furthermore, individuals with missing on any of the main explanatory variables female, foreign background and higher educated parent and educational aspirations and expectations have been removed from the analysis based on friendship network data. Table 1 shows the distribution of number of friends and figure 2 displays the characteristics of best friends.

[FIGURE 1 HERE.]

[FIGURE 2 HERE.]

3.4 Definitions of immigrant children

The analyses in this paper are based on two separate datasets with varied details on the birthplace of the respondents' parents. The advantage of the LNU data over the CILS4EU dataset is that it contains finer background variables. Using the LNU survey and register data it is possible to identify the region of birth of both parents. The LNU-UFB survey also contains information on the parents self-reported migration history, along with other useful background statistics. The CILS4EU dataset, however, only includes self-reported information by parents and children on whether the parents were born in or outside of Sweden.

In line with previous literature on immigration (Borjas 2011, Portes & MacLeod 1999), I use two commonly applied definitions of immigrant children. Throughout the main analysis of this paper I define children of immigrants as children with both parents born abroad regardless of own birthplace. The reference category is

---

[15]None of these "isolated" individuals filled in the Main questionnaire hence I am unable to explore their observable characteristics.

children with at least one parent born in Sweden. I refer to this category as *native-majority children.* I use the terms *immigrant children* and *children of immigrants* interchangeably. In the sensitivity analyses, I will also look separately at the outcomes of children who have themselves migrated. Ethnic groups are defined by parental birth region, as is common in the literature on ethnic differentials (see for example Szulkin & Jonsson (2007)).

In the analyses based on LNU data, the child's region of birth is based on the mother's region of birth. If there are two biological parents in the household, I use the mother's place of birth to define the child's ethnicity. If the mother was born in Sweden but the father is foreign-born, the father's region of birth defines the origin of the child. In case the household consists of a lone biological father, I use his region of birth.[16] I can distinguish eight regions of birth in the LNU-UFB dataset: *Sweden*, *Nordic countries*, *EU15+*, *Other European countries/Not EU15+*, *Middle East*, *Africa*, *Asia* and *Latin America* but due to a small sample size and multicollinearity I aggregate the regions into four categories: *Sweden*, *Europe* and *non-Europe*.[17]

3.5   Educational aspirations and expectations

I distinguish educational aspirations from educational expectations. Aspirations are defined as idealistic goals and refer to individual's hopes about the future regardless of constraints while expectations are what individuals think will happen when taking into account his or her constraints (Morgan 2006, Jacob & Wilder 2010).[18] The variables *educational aspirations* and *educational expectations* are drawn from the following two questions in the LNU-UFB 2010 questionnaire:

1. Would you like to continue going to school after the upper secondary level, that is, attend a university or university college? (Yes, absolutely; Yes, probably; No, probably not; No, absolutely not)

2. Do you think you actually will continue going to school after the upper secondary level, that is, attend a university or university college? (Yes, absolutely; Yes, probably; No, probably not; No, absolutely not)

---

[16]Using register data it is possible to distinguish between first and second-generation immigrants. I discuss alternative definitions in Appendix A.

[17]The variable indicating the region of birth of the partner of the interviewee in the LNU survey does not include the category *Middle East*. Children whose origin is based on the partner of the interviewee will have *Asia* and not *Middle East* as their region of birth.

[18]Educational expectations are dynamic, i.e. they are subject to constant revision and updating (Morgan 1998). The acquisition of new information on academic ability or on the costs and benefits of higher education may alter children's educational expectations (Jacob & Wilder 2010).

The outcome variable *aspirations-expectations gap* is defined as having aspirations for a university degree and but not expecting to get one. In the analysis that is based on LNU data the gap is coded "1" if an individual has responded "Yes, absolutely" or "Yes, probably" on the question "Would you like to continue going to school after the upper secondary level, that is, attend a university or university college?" and "No, probably not" or "No, absolutely not" on the question "Do you think you actually will continue going to school after the upper secondary level?"). The corresponding questions in the CILS4EU survey (first and second wave) are:

1. What is the highest level of education you wish to get? (Don't know, No degree, Compulsory school, Upper secondary school, College/university)

2. What is the highest level of education you think you will actually get? (Don't know, No degree, Compulsory school, Upper secondary school, College/university)

In the analysis utilizing CILS4EU data, the aspirations-expectations gap is coded "1" if an individual has responded " College/university" on question 1 above *and* less than university, i.e. any of the options "Don't know", "No degree", "Compulsory school", and "Upper secondary school" on question 2. In both samples, individual aspirations and expectations are measured close to a crucial transition point (from upper-secondary school to high school). The Swedish school system applies no formal tracking during the first nine years of compulsory school. At the end of middle school (last term in grade 9) students decide on their educational careers and what upper-secondary program to choose. The final year of grade 9 is a crucial transition point in the education system as students move from compulsory school to high school. GPA in ninth grade determines the set of feasible programs in high school. A pass in all core subjects is required for eligibility to secondary education.

3.6   Cognitive and language tests

The CILS4EU data include individual scores on both a cognitive and a language test. These two tests were administered in the first wave of the survey during the school year 2010-2011. The language test is a test of proficiency in the Swedish language. More precisely, it is a test of a child's lexicon of antonyms. The test includes 30 items with 4 alternatives each (for more information, see Kruse & Konstanze (2016)).

While the language test is used to assess children's verbal competencies, the cognitive test is "language free" and does not require any language skills. It is a 7 minute multiple-choice test of graphical puzzles including 27 items with properties similar to Raven's Progressive Matrices (Raven 2003). The maximum score of

this test is 27 and the minimum is 0. The distribution of test scores by parents' immigration status is shown in figures 3 and 4.

[FIGURE 3 HERE.]

[FIGURE 4 HERE.]

3.7   Explanatory variables

The explanatory variables are sorted into three categories: individual characteristics, family characteristics and friendship network variables. Descriptive statistics are presented in table 1.

[TABLE 1 HERE.]

I use a binary variable indicating the gender of the child (1=female). The age variable in the LNU 2010 dataset is the age of the student at the time of the survey (either 2010, 2011 or 2012). The average age of children in the LNU 2010 sample is approximately 16. Children in CILS4EU are in the age span 14-15 (wave 1) and 15-16 (wave 2). In the CILS4EU sample, I use a dummy of high skill occupation as a proxy for higher educated parents (1=if at least one parent has a high skill occupation). I use this variable rather than parents' self-reported years of education since less than 62 percent (n=3,104) of the respondents' parents participated in the survey and because the proportion of missing values is high among those that did take part.[19] In the analysis sample, about 43 percent are coded as managers and professionals.

In the analysis that is based on LNU data, I use the parents' self-reported level of education which should reduce the measurement error in the education variables. Parents' highest level of education is coded into three categories: compulsory school or less, high-school or less and university studies. Moreover, I distinguish three different types of families: intact families, reconstituted families and single-parent families. The proportion of children living in intact families is about 70 percent. I also use a variable indicating the number of individuals in the household drawn from survey answers.

School performance is measured by individual self-reported grades from the last school report in the core subjects Maths, Swedish and English. The response options are: Excellent, Pass with Distinction, Pass, and Fail and these are coded from 0 (Fail) to 3 (Excellent). These options are used to create a grade sum for each

---

[19]Engzell (2016) discusses the problems with using the parents' education in CILS4EU.

individual, ranging from 0 to 9.[20] The survey question regarding school grades is found in the second wave of the survey conducted during the respondents' final year in secondary school (compulsory school). Students' school grades in ninth grade will determine the set of attainable high school tracks (e.g. academic or vocational) and are thus crucial for their future academic career.[21] I use household characteristics (e.g. country of origin, education) from both surveys and Swedish registers. I standardize the achievement test scores (language and cognitive ability) for comparability and ease of interpretation.

Since I lack information on school performance for the LNU sample, I use students' self-assessed school performance, measured using the question: "If you compare yourself to your peers, how well do you think you do in school? (Best in class, Among the best, Better than the majority, About as good as most people, Not as good as most people)". The response options are coded as dummies.

I use each individual's self-reported friendship network (described in section 3.3 above) to calculate the average characteristics of the explanatory variables female, foreign background and higher educated parents and university aspirations among his or her best friends. I also calculate the corresponding means at the classroom and school level. Using the network information I create dummy variables for having best friends who all have the following characteristic: female, immigrant, higher educated parents and university aspirations. In addition to the variables listed above, I also use parents' aspirations and expectations of their child's education which are drawn from the following two questions found in CILS4EU:

1. What is the highest level of education you wish your child to get? (Don't know; No degree; Compulsory school; Upper secondary school; College/university)

2. What is the highest level of education think your child will actually get? (Don't know; No degree; Compulsory school; Upper secondary school; College/university)

Using these questionnaire items I create the variable *parent discrepancy* indicating whether the parent wishes the child to get college/university education but expects him or her to get less than college/university education.

---

[20] As a robustness test I also use indicator variables for Fail in each core subject and standardized grades in each subject. The results from these estimations display a similar picture (results are delivered upon request).

[21] The LNU survey also includes program in high-school (vocational and academic) and whether or not he or she has finished school.

4 Aspirations and expectations

4.1 Accounting for differences in aspirations and expectations

In this subsection, I examine the educational aspirations and expectations between children with two foreign-born parents and children with at least one native-born parent. Figure 5 reports the fraction of children with native versus foreign-born parents that answered each category on the question: "Would you like to continue going to school after the upper secondary level, that is, attend a university or university college?". This figure shows that children with foreign-born have answered "Yes, absolutely" to a larger extent than children with native-born parents (62.4 compared to 37.0 percent) and the difference is significant.[22] Children with foreign-born parents seem to show more "certainty" in their answers while children with native-born parents are more evenly distributed across the different categories.

[FIGURE 5 HERE.]

[FIGURE 6 HERE.]

Turning to figure 6 and the question: "Do you think you actually will continue going to school after the upper secondary level?", we see that children with foreign-born parents have replied "Yes, absolutely" to a larger extent than children with native-born parents (51.4 compared to 34.3 percent) and the difference is significant.[23] Again, children with foreign-born parents show more "certainty" in their answers, however the differences between the two groups are smaller compared to their differences in aspirations.

A first glance at the data shows that children of immigrants have both higher aspirations and expectations than their peers, but are they more or less consistent in their aspirations and expectations compared to the reference group? Contingency tables can be a useful way of measuring the inconsistencies between aspirations and expectations. The diagonal shows the proportion of individuals who answered consistently in tables 2 and 3. Full consistency corresponds to 100 percent in all the diagonal cells. Table 2 shows educational expectations by educational aspirations among children with native-born parents. This table shows how the children responded to the expectations question, given what they answered on the aspirations

---

[22]Conventional p<0.05 two-tailed t-tests of the differences between the other categories show that they are not significant.

[23]Conventional p<0.05 two-tailed t-tests of the differences between the other categories show that they are not significant

question. For example, given that they answered "No, absolutely not" on the aspirations question, 71 percent of the children with native-born parents answered No, absolutely not on the expectations question. The diagonal elements in table 2 are all larger than 70 percent for children with native-born parents.

[TABLE 2 HERE.]

[TABLE 3 HERE.]

Turning to children with foreign-born parents we see a similar pattern in table 3 where the diagonal elements are larger than 70 percent except for one particular case: among the children with foreign-born parents who answered "No, absolutely not" on the aspirations question, a significant proportion have replied "Yes, probably" on the expectations question. Note, however, that the sample contains 318 individuals with foreign-born parents and only a small fraction of the children answered "No, absolutely not" (as shown in figure 5).

Figure 7 reports childrens' responses to the question: "What is the highest level of education you wish to get?" and figure 8 show the distributions of the responses to the question: "What is the highest level of education you think you will actually get?". In line with the results presented above based on the LNU dataset, children of immigrants tend to have higher aspirations and expectations than their peers. Overall, children of immigrants do not show signs of higher or lower inconsistency than their native-majority peers.

Next, we turn to the CILS4EU dataset. Tables 7 and 8 present the aspirations and expectations of children in grade 8 by parents' immigration status. The results presented in the tables below validate the findings above: children of immigrants have both significantly higher aspirations (76.5 percent versus 62.4 percent, t-test α=5) and expectations (59.0 percent versus to 49.1 percent, t-test α=5) than children with native-born parents.

[FIGURE 7 HERE.]

[FIGURE 8 HERE.]

In sum, the analysis above seems to indicate that there is a significant difference in the aspirations and expectations among children with native-born parents and children with foreign-born parents. Children with foreign-born parents have, on average, both higher educational aspirations and expectations than their native-born counterparts before controlling for different individual and background characteristics. Next, I will try to account for the immigrant-native gap using rich sets of control variables found in the LNU and CILS4EU datasets.

The analysis is performed on two separate datasets. While the LNU dataset includes finer categories of the respondent's region of birth, it lacks information on school performance and social networks. I start by creating binary variables indicating aspirations and expectations for a university degree. The different sets of control variables in the LNU survey and matched administrative records fall into the categories: individual characteristics, family characteristics, and schooling variables. The unit of analysis is the child. I estimate the following linear probability model (LPM) by OLS:

$$y_i = \alpha + \beta_1 immstat_i + \beta_2 female_i + \beta_3 immstat_i \times female_i + \sum_{j=1}^{2} \beta_{4j} educ_{ji}$$
$$+ \sum_{j=1}^{4} \beta_{5j} selfass_{ji} + \sum_{j=1}^{3} \beta_{6j} famtype_{ji} + \beta_7 hhsize_i + \epsilon_i, \quad (1)$$

where y_i is a dummy indicating the outcome of child i in either educational aspirations, educational expectations or the aspirations-expectations gap, α is a constant, $immstat_i$ is a dummy indicating parents' immigration status (if the individual has two foreign-born parents and irrespective of own birthplace) and $female_i$ is a dummy representing gender. Standard errors are robust and clustered at the family level since the sample includes siblings. As the LNU data lacks information on individual school performance I instead include a measure of self-assessed school performance. This measure is drawn from the question "If you compare yourself to your peers, how well do you think you do in school?" with the response options: "Best in class", "Among the best" and "Better than the majority", "About as good as most people", and "Not as good as most people". The term $selfass_{ij}$ denotes indicators of self-assessed performance. The reference category consists of individuals who answered "Not as good as most people". $educ_i j$ denotes a set of indicators of the highest education of both parents (interviewee and partner) and the categories include: compulsory school or less, high-school or less and university studies. Observations are coded as missing if both the interviewee and partner have missing values on the relevant variable. $famtype_i j$ stands for a set of covariates representing either intact family, reconstituted family, single parent household or other. Each response category of the family type question has been coded as a dummy variable. $hhsize_i$ is a continuous variable indicating the number of members in the household.

In table 4 I present the results from OLS regressions on the three outcome variables listed above. The models are estimated for a sample of children in the ages 12-19.[24] The coefficient of interest is β_1 which captures the discrete change

---

[24]Alternative model specifications can be found in the sensitivity analysis in Appendix A. Table A1 presents both the log odds and average marginal effects.

from having at least one native-born parent to having two foreign-born parents. Children with immigrant parents have significantly higher aspirations than their native majority peers; they are 17.8 percentage points (ppt) more likely to report having aspirations for a university degree.

[TABLE 4 HERE.]

In model (2) I enter all the covariates listed in section 3.7 hence β_1 is now interpreted as the independent effect of parents' immigration status once all the factors above have been controlled for. The sample size is only marginally reduced due to missing values on the covariates but the magnitude is unchanged implying that the results are robust to including individual and family characteristics.

To assess whether the relationship differs by gender, I also include an interaction term between the indicators for having two foreign-born parents and being a female. Due to the inclusion of the interaction term, the reference category is now native-majority boys. The interaction term indicates by how much the influence of being an immigrant student differs between girls and boys and ceteris paribus, the "net effect" of having two immigrant parents is less positive for girls than boys as the interaction coefficient β_1 is -17.1. Interestingly, the results display a significant gender gap: all else equal, immigrant girls are 30 ppt more likely to have aspirations for a university degree than native-majority boys. An F-test indicates that the gender differential between immigrant boys and immigrant girls is significant (α=0.05, column (2)). Moreover, they are 5 ppt more likely to express university aspirations than immigrant boys, although this difference is insignificant (F-test, α=0.05, column (2)). Native-majority girls are significantly more likely than native-majority boys to report university aspirations (21.6 ppt, p<0.01). Girls with two foreign-born parents are 8.4 ppt more likely to have university aspirations than native-majority girls (not significant).

Unsurprisingly, children with higher educated parents are more likely to wish to study at university than their peers (18.0 ppt, p<0.01). Self-assessments of own scholastic ability seems to matter too. Individuals who consider themselves belonging to the upper part of the ability distribution among their peers tend to report university aspirations to a larger extent (ranging between 39.6 to 22.7 ppt, p<0.01) than the reference group consisting of individuals who have responded "Not as good as most people". The positive influence seems to grow with higher self-assessment.

In model 3 (as well as in model (6) and (9)) the binary outcome variable $immstat_i$ is replaced by three dummy variables indicating region of origin: "Sweden", "Europe" and "non-Europe".[25] Moreover, the results are shown conditional on gender

---

[25]Figure D.1 in Appendix D.1 shows the distribution of region of origin in the LNU sample.

and socioeconomic background. The β_1 coefficient now indicates the differential between children with an European or non-European origin and children with at least one native-born parent. Children with parents of European origin are not significantly different in their university aspirations than their native-majority peers. Children with parents born in a non-European country are, however, 21.4 ppt more likely to have university aspirations than their peers (p<0.01).

Moving on to educational expectations and model (4), β_1 is the raw difference in educational expectations between children with two immigrant parents and children with at least one parent born in Sweden with the same gender, similar socioeconomic and family background, and self-assessments of scholastic performance. The immigrant-non-immigrant gap is also positive but smaller when we look at the educational expectations in model (4) compared to aspirations in columns (1) and (2). The probability of thinking you will attend university is 15.7 ppt higher among children with foreign-born parents than the reference category (p<0.01). The estimate increases in the analysis sample (15.7 ppt versus 22.7 ppt). The direction and relevance of the estimates are similar to those in the models of aspirations (models (1)-(3)). The gender differential is present also with respect to university expectations. Both immigrant and native-majority girls are significantly more likely to report expectations for a university degree (28.7 ppt and 21.7 ppt).

The immigrant-non-immigrant differential is significant and larger for children with an non-European origin (around 18.6 ppt, p<0.01) compared to children of European decent. Both non-European girls and boys have higher aspirations than their native-majority counterparts (not shown here). University expectations are more likely among girls than native-majority boys (21.7 ppt higher among native-majority girls and 28.7 ppt higher among immigrant girls). In addition, the gap in aspirations between non-European boys and native-born boys is larger than the gap between non-European girls and native-born girls (not shown here). Family type and household size do not seem to matter for aspirations and expectations (reference category is intact family and other).

Moving on to models (7)-(9) the dependent variable is the aspirations-expectations gap which is defined as having aspirations for a university but not expecting to get one. The β_1 coefficient is close to zero and insignificant. The proportion of children expressing a gap is only 2.6 percent (displayed in the descriptive table in section 1) which explains the low explanatory power of the model. I find no significant differences in the immigrant-non-immigrant aspirations-expectations gap presented in model (7), (8) and (9), suggesting that immigrant children are not more likely to have expectations falling short of aspirations than their native-majority peers. Model (9) also shows that, all else equal, native-majority girls are significantly less likely to express a gap compared to their male native-majority

peers. Due to the small sample size these results need, however, to be interpreted with caution.

In sum, children with foreign-born parents seem to have both higher educational aspirations and expectations than their counterparts with native-born parents. After controlling for parents' highest education there is still a significant gap in both aspirations and expectations. An important piece of the puzzle is still missing – academic potential and performance – factors that will be added in the subsequent analysis of section 4.2. Perhaps the remaining gap could be explained by some subgroups having unrealistically high aspirations and expectations with respect to their academic performance or potential? Another possible explanation is unobserved heterogeneity among students such as behavioral factors, motivation, parental preferences towards higher education.

In the next step I broaden the analysis by using the CILS4EU dataset which includes self-reported school results and friendship links. As before, I try to account for differences in aspirations and expectations among children with foreign-born and native-born parents but this time *conditional* both on academic potential and school performance. I then add sociometric information and different socio-economic, schooling factors found in CILS4EU. Due to possible sorting across schools, in what follows I will include classroom fixed-effects in all of the regression models.

4.2 The role of school performance and friendship network

Previous research shows that foreign-born children and children of immigrants lag behind children of native-born parents in educational performance in several European countries, and Sweden is no exception (Schnepf 2007).[26] Foreign-born students in Sweden are less likely to be eligible to attend upper secondary school than their native-born counterparts.[27] Importantly, immigrant children is a heterogeneous group. Given attained school grades they tend to make more ambitious study choices than their native-majority peers (see Arai et al. (2000), (Jonsson & Rudolphi 2011) and Heath & Brinbaum (2014)). The school performance of children born in Sweden with immigrant parents is varying by ethnic origin and those who decide to continue to higher eduction studies often outperform the majority population, while those who fail in secondary school have low labor market prospects (Jonsson & Rudolphi 2011).

Figure 11 shows the distributions of grades by parents' immigration status (ir-

---

[26]See also Heath & Brinbaum (2007) and Heath & Brinbaum (2014).

[27]Arai et al. (2000) show that children born abroad have lower marks, attain lower levels of education and show higher risks of unemployment than their native-born counterparts.

respective of own birthplace).[28] Children with immigrant parents are more likely to have non-complete grades in core subjects. The largest difference between the two groups is in English and the second largest difference is in Maths where 9.3 percent of the sample of children with immigrant parents have reported Fail on their last school report (versus 5.7 percent among their native-majority peers).

[FIGURE 11 HERE.]

Next, I turn to the CILS4EU dataset. This dataset includes friendship network data and extensive survey information on parents and friends. I start by estimating the following regression:

$$
\begin{aligned}
y_{ict_0} = \alpha + \beta_1 immstat_i + \beta_2 female_i + \beta_3 immstat_i \times female_i + \beta_5 educ_i \\
+ \sum_{j=1}^{2} \beta_{4j} testscore_{ji} + \beta_6 pardisc_i + \mathbf{x}_{it_0}^{friend} \boldsymbol{\gamma} + \epsilon_{ict_1},
\end{aligned}
\tag{2}
$$

where y_{ict} is a dummy indicating the outcome variable either university aspirations or university expectations of individual i in grade 8 denoted by t_0 and class c. α is a constant, $immstat_i$ is a dummy indicating if the individuals has two foreign-born parents irrespective of own birthplace and $testscore_{ji}$ denotes individual test scores on cognitive and language tests performed in grade 8 (wave 1). Parental level of education is proxied by parents' occupational class and $educ_i$ is a dummy denoting high skill occupations and coded 1 if at least one parent has an occupation belonging to the categories managers and professionals.[29] Parent discrepancy, represented by the term $pardisc_i$, is a dummy indicating whether the parent wishes the child to get college/university education but expects him or her to get less than college/university education, and ϵ_{ict_0} is the error term.

The characteristics of friends, denoted by $\mathbf{x}_{it_0}^{friend}$, are four dummy variables indicating whether or not an individual has friends who all have the following characteristics: female, foreign-born parents, higher educated parents and university

---

[28]Children who have themselves immigrated are included in the sample of children with immigrant parents which could explain a large part of the immigrant-non-immigrant differential displayed in figure 11. In the regression analysis that follows, I control for language proficiency, among other things.

[29]The dummy managers and professionals is coded 1 if either parent's occupation belongs to the categories >999 & <3000 of ISCO 2008. Parents' occupational class is drawn from the children's reports of their parents' occupation. The variable is drawn from the student questionnaire item: "Think about your mothers job. If she is not currently working, think about her last job. What is the name of her job? Additionally, please describe what she does in her job.". I also use the corresponding question for father's occupation. An alternative approach is to create three categories based on ISCO 2008: "high skill": 1 and 2, "mid skill": 3,4 and 7 and "low skill": 5,6,8 and 9.

aspirations.[30]

Due to potential sorting across classrooms and schools as well as different grading practices across schools I include classroom fixed-effects in all of the models (N.B. I use the terms school class and classroom interchangeably). Moreover, by including classroom fixed-effects I control for the educational environment such as teacher and classmate quality. The classroom dummies absorb across classroom and school differences, i.e. I only compare students in the same classroom. FE_{ct_0} denotes fixed-effects at the classroom level and the indicator is based on the composition of students on the day of the network survey in grade 8. Unless a student relocates or actively decides to change class, students in Sweden tend to have the same classmates all through grades 7-9 of compulsory school.

Next, I match the first wave of CILS4EU with the second wave that also contains self-reported grades in Maths, Swedish and English in grade 9 in last semester (not final grades). Thus, I am able to investigate differences in aspirations and expectations between children of foreign-born and children of native-born *conditional* on academic performance. I proceed by estimating the following regression:

$$y_{ict_1} = \alpha + \beta_1 immstat_i + \beta_2 female_i + \beta_3 immstat_i \times female_i + \beta_4 educ_i$$
$$+ \sum_{j=1}^{2} \beta_{5j} testscore_{ji} + \beta_6 pardisc_i + \beta_7 gradesum_i \qquad (3)$$
$$+ \mathbf{x}_{it_0}^{friend} \boldsymbol{\gamma} + \nu FE_c + \epsilon_{ict_1},$$

where y_{it} is a dummy indicating either the outcome variable university aspirations or university expectations of individual i in grade 9, class c, and wave 2 as denoted by t_1. The dependent variable is a dummy defined as having aspirations for a university degree, i.e. having reported "College/university" on the questionnaire item "What is the highest level of education you wish to get?". The terms α, $immstat_i$, $educ_i$, $testscore_{ji}$ and $pardisc_i$ are defined as above. As before, $\mathbf{x}_{it_0}$ represents a vector of covariates at the friendship level where friends are defined according to friendship nominations at t_0 (wave 1). Since I only use the friendship data in t_0 I treat the friendship network as static. The term $grade_{jit_1}$ represents an individual's grade in Maths, English and Swedish. To correct for potential similarity of individuals within clusters the standard errors are clustered at the classroom level. Table 5 reports OLS regression coefficients from models of university aspirations.[31]

---

[30]In alternative specifications the vector includes the mean characteristics of best friends. Moreover, I also include the average peer characteristics at the classroom and school level, represented by the vectors $\bar{\mathbf{x}}_{it_0}^{class}$ and $\bar{\mathbf{x}}_{it_0}^{school}$. Alternative specification results are either presented in Appendix A or available upon request).

[31]I also run multileveled logit regressions with random-effects at the classroom level and standard errors clustered at the classroom level. These models are estimated using Stata's melogit command

In table 5 I present the estimations of the outcome variables aspirations in grade 8 (models (1)-(3)) and aspirations in grade 9 (models (4)-(7)). Recall that $immstat_i$ is a dummy hence β_1 captures the change in likelihood as the variable changes from 0 to 1. Model (1) shows the raw difference in aspirations between children with at least one native-born parent and children with two foreign-born parents in grade 8 for the full sample. Immigrant children are 10.7 ppt (p<0.01) more likely to express university aspirations than their native-majority peers.

Model (2) shows the corresponding results for the analysis sample. The coefficient is only slightly higher (10.7 versus 11.7) indicating that the result is robust to sample changes. Unconditional on individual, friends and family characteristics, children of immigrant background are significantly more likely to want to study at university than their peers.

In model (3) I partial out the influence of gender, parents' highest level of education, academic potential (proxied by cognitive and language ability tests), parents' aspirations-expectations discrepancy and the average characteristics of friends. The characteristics of friends include: proportion of female friends, proportion of friends with two foreign-born parents and proportion of university aspiring friends. The coefficient for the dummy indicating immigration status indicates that immigrant boys are 25.7 ppt more likely to have university aspirations than native-majority boys. The interaction term indicates by how much the influence of being an immigrant student differs between girls and boys. The influence of immigration status is less positive for girls compared to boys but not significantly (F-test, α=0.05, column (3)). Native-majority girls are significantly more likely to report university aspirations than native-majority boys (approximately 10.8 ppt more likely, p<0.01). Girls with foreign-born parents are 28.9 ppt more likely to express university aspirations compared to the base group (F-test, α=0.05, column (3)).[32] Thus, the gender difference seems to be larger within the native-majority group than the immigrant group.

Individuals with at least one higher educated parent (proxied by both parents' occupational class) are significantly more likely to wish to attend university (10.5 ppt more likely). Language proficiency seems to matter more than cognitive ability: all else equal, a standard deviation (std=4.85) increase in the language test score of an individual is on average associated with a 11.6 ppt increase in the probability of

---

for mixed multilevel logit models.

[32]The coefficient for girls with foreign-born parents is: 0.257+0.108-0.0756=0.289.

having university aspirations. The estimate of the influence of the cognitive ability test score is also significant but smaller: 4.0 ppt (std=4.72).

To determine the importance of the composition of an individual's social networks I also include a selection of observable characteristics of an individual's friendship network. Having only female friends is negatively associated with university aspirations. However, the coefficient is small and insignificant.[33] Other friends' characteristics seem to be positively correlated with individual aspirations, for instance the estimate is positive for the indicators for having only friends with foreign-born parents (4.41 ppt) and only friends with university aspirations (1.52). All estimates are insignificant at conventional levels suggesting that friendship characteristics are not major determinants of aspirations in grade 8. A reason for this could be that the fixed-effects at the classroom level have captured much of the variance.[34]

Next, I present the results for the same individuals only now their aspirations and expectations have been measured one year later in grade 9. Model (4) shows the raw difference between children with immigrant parents and at least one native-born parent in the same classroom in the full sample (9.46 ppt). The analysis sample in model (5) consists of 3,381 individuals. The estimate of having two foreign-born parents increases somewhat to 13.3 ppt indicating some degree of selection from the reduced sample size.

Moving on to model (6), we can see that the differential is higher when controls are entered (23.5 ppt, p<0.01). The classroom fixed-effects imply that I am comparing individuals in the same classroom with similar individual characteristics, academic potential, school performance and social background.[35] Model (6) also holds constant for grade sum measured by adding up the grades in each core subject thus I adjust for the presence of individuals with unrealistic aspirations and expectations. Given the same academic potential and similar school results, children of immigrants have significantly higher aspirations than their native-majority peers.[36]

Language test scores are still important. The association between aspirations and language tests scores is weaker in grade 9 than in grade 8 but still significant compared to the previous influence of cognitive ability which disappears (11.6 ppt versus 3.1 ppt). Moreover, grade sum in the core subjects Maths, English and Swedish are positively related to university aspirations: all else equal, a one point increase in the grade sum (moving from a "Pass" to a "Pass with Distinction" in a

---

[33]In alternative specifications I have used the proportion of friends with certain observable characteristics. These estimates are both very small and statistically insignificant.

[34]The results from fixed-effects logistic regressions in Appendix A show a similar picture.

[35]In alternative specifications of model (6) we instead hold constant for having passed all core subjects (based on self-reported grades from last school report in grade 9).

[36]As previously mentioned, the grades are based on self-reported grades from the last school report in grade 9.

core subject) is associated with 7.2 ppt increase in the likelihood of having university aspirations (p<0.01).

The results from regressions including friends characteristics are presented in model (7). Overall, the estimates have the same direction as in model (3) and (6). The results suggest that the influence of the observable characteristics of friends are all insignificant. University aspirations among boys with two foreign-born parents is now 23.4 ppt (p<0.01) higher than that of the base category which consists of native-majority boys. Furthermore, native-majority girls and immigrant girls are more likely to report university aspirations than is the base category: the coefficient for immigrant girls is 26.8 ppt (F-test, α=0.05, column (6)) and for native-majority girls it is 11.5 ppt. Holding constant for the characteristics of an individual's social network, the immigrant-non-immigrant differential is still present and significant.

Following the same procedure as above I estimate the corresponding regressions for the outcome variable educational expectations and the results are reported in table 6. The dependent variable is a dummy defined as having expectations for a university degree, i.e. having reported "College/university" on the questionnaire item "What is the highest level of education think you will actually get?".

Model (1) in table 6 presents the raw difference in university expectations between children of immigrants and native-born children for the full sample in grade 8. The coefficient of particular interest β_1 indicates that immigrant children are 7.14 ppt more likely to want to attend university than their native-majority peers (p<0.01). The coefficient is larger and still highly significant in the analysis sample in model (2) (8.02 ppt versus 7.14 ppt).

Model (3) shows the results conditional on the set of explanatory variables in the first wave of CILS4EU. The signs and relevance are similar to the models of university aspirations in table 5. All else equal, models (1)-(3) suggest that children of immigrants report expectations for higher education in grade 8 to a larger extent than their native-majority peer.

The outcome variable in models (4)-(7) is educational aspirations in grade 9. Similarly to the model of expectations in grade 8, the estimate for having immigrant parents increases when covariates are added into the model of expectations in grade 9 (13.5 ppt versus 24.6 ppt). Unlike the cognitive ability test scores which looses significance in grade 9 expectation models, language ability remains highly significant in both model (6) and (7).

I enter the full set of explanatory variables in model (7) and now I am comparing children in the same class with similar individual and family characteristics, school performance and academic potential (proxied by language and cognitive tests in grade 8) and friendship network characteristics. As expected, the grade sum in core subjects is positively related to expectations for higher education (model (7)).

With regard to the friendship level variables, all except having only friends with higher educated parents turn out to be insignificantly associated with individual expectations (p<0.01).

To compare the fit of the network model and the full-sample model I perform a log-likelihood test (of whether the models are nested) where the null hypothesis states that there is no significant difference between the two models. The test indicates that I can reject the null hypothesis and conclude that the network model provides a better fit to the data.

To summarize the tables for aspirations and expectations, children of immigrant parents are more likely to aspire to university studies and to expect to study at university than their native-majority peers. Both immigrant and native-majority girls are more likely to aspire to and expect university studies than native-majority boys (F-tests, $\alpha=0.05$, columns (6)-(7)).

5   The aspirations-expectations gap

In this subsection, I examine whether children of immigrants are less consistent in their aspirations and expectations than their native-majority peers. I treat the outcome variables aspirations and expectations as ordinal since the difference between "Yes, absolutely" and "Yes, probably" is not necessarily the same as the difference between "No, probably not" and "No, absolutely not". Furthermore, the error terms of the two models may be correlated. Simply subtracting the one from the other will be inappropriate which is why I consider the following two alternative approaches.

The first approach involves dichotomizing the underlying variables aspirations and expectations and running multileveled logit regressions on a binary outcome variable defined as having aspirations for a university degree but not expecting to get one. A drawback with this method is that is could lead to a loss of important information from the model. The alternative strategy is to estimate the model using the bivariate ordered probit model thus making full use of the categorical property of the aspirations and expectations variables in the LNU dataset. The latter approach is less suitable for the CILS4EU data and for comparability the former is therefore more appropriate. Hence, in what follows I dichotomize the aspirations-expectations gap. I present the results from the alternative approach in Appendix A. I follow the

same procedure as in section 4.2 above and estimate the following regression:

$$y_{ict} = \alpha + \beta_1 immstat_i + \beta_2 female_i + \beta_3 immstat_i \times female_i + \beta_4 educ_i$$

$$+ \sum_{j=1}^{2} \beta_{3j} testscore_{ji} + \beta_5 pardisc_i + \beta_6 gradesum_i + \mathbf{x}_{it_0}^{friend} \boldsymbol{\gamma} \quad (4)$$

$$+ \nu FE_{ct_0} + \epsilon_{ict},$$

y_{ict} is the aspirations-expectations gap of individual i in t_0 (wave 1) or t_1 (wave 2). The outcome variable y_{it} is a dummy indicating whether an individual has aspirations for a university degree and but does not expect to get one, i.e. having responded "College/university" on the question "What is the highest level of education you wish to get?" and having replied any of the options "Don't know", "No degree", "Compulsory school" or "Upper secondary school" on the question "What is the highest level of education you think you will actually get?". All other variables are defined as above. The OLS regression results are presented in table 7. Models (1) and (2) report the baseline regression results of the outcome variable aspirations-expectations gap in grade 8 on the dummy $immstat_i$ including class fixed-effects. Standard errors are clustered at the classroom level.

[TABLE 7 HERE.]

The estimate is positive and significant (p<0.05). Children who have two foreign-born parents are 3.12 ppt more likely to express a gap than the native-majority group (the unconditional average of the aspirations-expectations gap among native-majority children is 0.167). The estimation results also reveal that the positive association between parents' immigration status and the aspirations-expectations gap disappears once covariates are entered into the model (table 7, model (3)). The adjusted R^2 is increases but only marginally.

Gender seems to play a non-negligible role since the coefficient for gender is bordering on statistical significance (p<0.10). Native-majority girls are 2.72 ppt more likely to express a gap than are native-majority boys, all else equal. Immigrant girls are also more likely to have mismatched aspirations and expectations than native-majority boys (4.28 ppt, F-test, α=0.05, column (3)). However, the gender differential is not significant within the immigrant group (1.86 ppt, F-test, α=0.05, column (3)).

In models (4)-(7) we turn to the outcome variable aspirations-expectations gap in grade 9. Interestingly, the gap is no longer present in the baseline regressions of models (4) and (5) suggesting that individuals adjust their aspirations over time. Children with higher educated parents are less likely to report an aspirations-

expectations gap (4.0-4.1 ppt, models (6) and (7)). Moreover, grade sum is negatively associated with reporting a gap.

Due to the low response rate among parents the results from regressions including the parent discrepancy indicator are presented in a separate table. Table 8 suggest that parent discrepancy is an important predictor of both aspirations and expectations. Having a parent with expectations that fall short of aspirations is negatively related to individual aspirations and expectations for higher education (p<0.01). The coefficient stays negative and significant in all specifications. Children's aspirations-expectations gaps are also significantly associated with their parent's aspirations as demonstrated in the two far right columns of table 8.

[TABLE 8 HERE.]

To summarize table 7, children of immigrants are significantly more likely to express an aspirations-expectations gap in grade 8. However, the gap diminishes over time and once covariates are entered into the model. Overall, neither immigrant boys nor immigrant girls are more likely to have mismatched aspirations and expectations than are their native-majority peers. The results validate the findings in section 4.1. The next question is whether individuals with non-complete grades are more likely to have unrealistically high aspirations and if children of immigrants are more likely to belong to this student category.

5.1   Sub-sample analysis

In this subsection, I look closer at the individuals who have non-complete grades. To examine potential subgroup differences, I run separate analyses on individuals with complete and incomplete grades. I create a non-complete grades dummy that is based on self-reported grades from the last school report (1=at least one Fail in any of the core subjects) which is used as a proxy for school performance. The reference category consists of individuals with at least Pass in all core subjects. Figure 11 reports the self-reported grades in the core subjects Maths, Swedish and English. Distributions of school performance are shown in figure 12 below.

[FIGURE 12 HERE.]

Next, I investigate whether individuals with non-complete grades are more likely to have unrealistically high aspirations in grade 9. Table 9, Panel A, shows that within the subgroup of individuals with incomplete grades, children of immigrants do not differ significantly from the native-majority children. The coefficients for the immigration status of parents in column (1) and gender in columns (2)-(3) are

both positive but none of the estimates is significant. While column (1) reports the "raw" differential including fixed effects, the estimated models in columns (2) and (3) account for individual and social network characteristics.

[TABLE 9 HERE.]

Turning to Panel B and the subgroup of students with at least Pass in all core subjects, boys of foreign background are approximately 3.0 ppt (column (2)) less likely to have mismatched aspirations and expectations than boys of native-majority background, conditional on parental education and language proficiency. This estimate is, however, insignificant. The set of explanatory variables is expanded in column (3). In addition to the variables listed above, the regression model in column (3) also comprises of friends' characteristics including foreign background, female, higher educated parents. Among children with complete grades, those with at least one higher educated parent are 4.21 ppt less likely to express an aspirations-expectations gap ($p<0.01$, column (3)). Language test scores are significantly and negatively associated with reporting an aspirations-expectations gap in grade 9.

6 "Lost talent" among immigrant youths?

Following Hanson (1994) I define lost talent as scoring higher than the mean of the sample on the cognitive ability test administered in grade 8 and expressing an aspirations-expectations gap. The cutoff is produced using the analysis sample in previous sections (n=4,075). The distribution is skewed to the left (skewness is negative) and the median is greater than the mean (19.0 versus 17.9). The minimum score on the cognitive ability test for the analysis sample is 0 and the maximum is 27. In column (1) the threshold is set at a test score higher than the mean score in the analysis sample (henceforth referred to as definition A) while in column (2), lost talent is defined as scoring higher than the median in the analysis sample (henceforth referred to as definition B). According to definition A, 9.99 percent of the sample irrespective of parental migration status are labeled as lost talent in grade 8 and 6.55 according to definition B. In grade 9, the proportions are lower: 5.77 and 3.73 percent respectively.

Table 10 shows OLS coefficients for models predicting lost talent among a representative sample of eight and ninth graders in Sweden. All models have very low adjusted R^2-levels. I leave out the fixed-effects to keep more variation, i.e. I do not confine the analysis to variation within classrooms since the number of individuals labeled as lost talent per classroom is too small.

Column (1) reports that according to definition A, children with immigrant parents are 1.7 ppt less likely to belong to the category of lost talent ($p<0.10$). Covariates are added into the model in column (2) and the results indicate that neither boys nor girls with immigrant parents are significantly more likely to be labeled as lost talent than the reference group (native-majority boys).[37] All else equal, native-majority girls are significantly more likely to belong to the lost talent category than are native-majority boys (2.61 ppt, $p<0.05$, column (2)). The result is, however, not robust with respect to alternative definitions (A and B and in grade 8 and 9).

[TABLE 10 HERE.]

According to definition B and unconditioned on covariates, there is a significant immigrant-non immigrant differential in lost talent (1.89 ppt, $p<0.05$, column (3)). Thus, native-majority students are more likely to be classified as lost talent. The gap diminishes, however, as covariates are added into the model (column (4)). Columns (5)-(8) reveal a similar pattern. Overall, the immigrant-non-immigrant disparity is negative and insignificant. Language proficiency is a significant and consistent predictor of lost talent. Moreover, having only university aspiring best friends seems to be important too, although the estimate is only significant in column (8) (1.34 ppt, $p<0.05$). Apart from having only university aspiring friends there is a another network indicator that turns out to have a significant influence on the outcome variable: having only female friends is negatively associated with lost talent both in model (2) and (3). The estimated association is approximately 2 ppt ($p<0.10$ in column(2) and $p<0.05$ in column (3)).

To summarize table 10, the results suggest that immigrant children align their aspirations and expectations according to their school results. Conditional on school performance and academic potential, I find no indication of immigrant children being more likely to be labeled as lost talent in Sweden. I find that native-majority girls are more at risk of being labeled as lost talent in grade 8 than are their male native-majority peers according to definition A. The results is, however, not robust to changing cutoff definition.

7   Discussion

The issue of educational aspirations and expectations is important from a policy perspective since expectations are strong predictors of educational attainment (see for example Feliciano & Rumbaut (2005), Jacob & Wilder (2010), Morgan (2005),

---

[37]Immigrant girls are 1.61 ppt more likely to be labeled as lost talent according to column (2) but we cannot reject the null from an F-test that the means are the same ($\alpha=0.05$).

Portes & Rumbaut (2001)), which in turn is key for the economic and social integration of children with immigrant background (Bratsberg et al. 2011, Card 2005, Card & Schmidt 2003).

In this study, I explore a potential mechanisms for immigrant-non-immigrant disparities in school results, namely the aspirations-expectations gap. In line with previous research (Guyon & Huillery 2016, Salikutluk 2016, Hanson 1994, Rudolphi 2014, Heath & Brinbaum 2007) and based on two independent and nationally representative samples, I find that children of foreign-born parents tend to have higher aspirations and expectations than their native-majority peers. Conditional on a set of background factors, individual academic performance and academic potential, they are on average 25-30 ppt more likely to have university aspirations and express university expectations than their peers. One interpretation of these findings is that children of immigrants value higher education more than their native-majority counterparts: many seem to be well-aware of the importance of education for moving up the social ladder.

Part of the differential could also be explained by measurement error in the explanatory variable parental education proxied by occupational status. Immigrants are more likely to be overeducated (Joona et al. 2014) and the variable used in this study is likely not fully capturing the relative social class of immigrant children. A handful of recent studies have investigated this issue further (e.g. Ichou (2014) and Engzell (2016)) by testing whether a better measure would be to use the social class in the country of origin as some immigrant groups are positively selected on both observables (e.g. educational attainment) and unobservables (e.g. "motivation" and "drive").[38] Thus, an explanation behind the relatively high aspirations and expectations of immigrant students could be that they aspire to or expect to attain the social status of their parents' in the country of origin. Another potential mechanism behind the immigrant-non-immigrant disparities in aspirations, expectations and educational choice could be a strategy to avoid discrimination: immigrant students may aim higher because they want reduce the risk of discrimination in low-skill occupations (Rudolphi 2014).[39]

Overall, I do not find evidence of a significant aspirations-expectations gap among immigrant children: their aspirations and expectations are not more likely to be unaligned than their native-majority peers'. The results suggest that the gap is not the main driving mechanism behind immigrant-non-immigrant disparities in school outcomes.

In contrast to the findings of Rudolphi (2014) and Salikutluk (2016), I find that

---

[38]See discussion in for example Ichou (2014) and Engzell (2016).

[39]See Carlsson & Rooth (2007) on ethnic labor market discrimination by occupation.

both immigrant and native-majority girls are significantly more likely to express a gap in grade 8 than are native-majority boys. The results also display a significant gender differential in the outcome variable lost talent, a category defined as students scoring higher than the mean or median of the sample on the cognitive ability test and who are reporting an aspirations-expectations gap. I find that native-majority girls are significantly more likely to be labeled as lost talent. The gender and socioeconomic composition of an individual's closest friends seems to matter too. Having only female friends is negatively related to showing signs of academic potential and expressing an aspiration-expectations gap. In line with Hanson (1994), I find that the educational aspirations (or educational values) of friends is an important predictor of lost talent.[40]

Clearly, the result from the lost talent analysis hinges on the definition of lost talent which concerns two main factors: (i) the relevance of the threshold (ii) whether the variable cognitive test scores is an appropriate indicator of school talent or academic potential.

Based on these findings, an important avenue for future research is the role of teachers' and parents' aspirations and expectations for student outcomes, such as for instance the decision to dropout. Among those who decide to dropout, the lion's share are children of immigrant background. Another important policy question is whether aspirations and expectations are set efficiently. For example Guyon & Huillery (2016) find evidence of significant biases in aspirations among low-SES students who tend to aspire lower and are more likely to have fatalistic views. Two categories of students are therefore of particular interest from a policy perspective: low aspiring but high-achieving low SES-students and high-aspiring but low-achieving low SES-students. If aspirations malleable adjusting them through tutoring or parent-teacher meeting could bring about substantial improvements in students' outcomes at a relatively low cost.

As of today, causal studies of the role of individual aspirations for school outcomes are scarce, yet the field is growing (Goux et al. forthcoming, Avvisati et al. 2014, Carlana et al. 2015).[41] The existing evidence suggest that teachers and parents aspirations play an important role in the formation of individual aspirations and in educational choice. It is argued that setting realistic aspirations can have significant effects on individual life outcomes and that interventions aimed at low-performing but high aspiring or high-performing low-aspiring students could be relatively in-

---

[40]The definition of the lost talent is somewhat different in Hanson (1994) compared to this study which limits comparability between the two studies. The reference category in Hanson (1994) consists of students who both aspire to and expect to attain a college degree. Moreover, the study of Hanson (1994) is based on a sample of American high school seniors.

[41]See a detailed review of the literature in Fryer Jr (2016).

expensive compared to for instance reduced classroom size. An important question is: whom should be targeted? I find that parents aspirations play an important role in children's educational plans. Mismatched parental expectations, i.e. having a parent who is expressing an aspirations-expectations gap, is both a statistically and economically significant predictor of the individual aspirations-expectations gap. However, a widely recognized issue with targeting parents is selection, i.e. those who need the educational program the most are more likely to opt out (see for example Goux et al. (forthcoming)) and as an additional point longterm interventions are expensive and cost-efficiency is questionable in some regards.

One way to empirically test the importance of the aspirations-expectations gap is to track those identified as "lost talent" as they move through the educational system. Are they more likely to dropout after comprehensive school? What is their highest level of attained education as adults? Such an approach is possible by matching the data used in this study data with comprehensive administrative data on students' educational outcomes.

A potential issue with using survey data is individual non-response. Those who took part in the survey are perhaps more likely to have high educational aspirations and value education higher than the absentees. It is not unlikely that the respondents are positively selected on these characteristics. If immigrant students with low aspirations are more likely to shirk, the estimates of the immigrant-non-immigrant gap will be exaggerated.

To sum up, dropouts and students with incomplete grades is a highly ranked issue on the political agenda. Understanding the mechanisms underlying the decision to dropout is essential for effective policy formation. Children's aspirations and expectations are not formed in a vacuum and the role of teachers' and parents' aspirations for students' outcomes are two important avenues for future research.

References

Akerlof, G. A. & Kranton, R. E. (2002), 'Identity and schooling: Some lessons for the economics of education', *Journal of Economic Literature* **40**(4), 1167–1201.

Alm, S. (2011), 'The worried, the competitive and the indifferentapproaches to the future in youth, their structural roots and outcomes in adult life', *Futures* **43**(5), 552–562.

Appadurai, A. (2004), The capacity to aspire: Culture and the terms of recognition, *in* V. Rao & M. Wallton, eds, 'In Culture and Public Action', Stanford University Press.

Arai, M., Schröder, L. & Vilhelmsson, R. (2000), *En svartvit arbetsmarknad: en ESO-rapport om vägen från skola till arbete: rapport till Expertgruppen för studier i offentlig ekonomi*, Fritzes offentliga publikationer.

Avvisati, F., Gurgand, M., Guyon, N. & Maurin, E. (2014), 'Getting parents involved: A field experiment in deprived schools', *The Review of Economic Studies* **81**(1), 57–83.

Becker, G. S. (1964), *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, The University of Chicago Press.

Björklund, A., Edin, P.-A., Holmlund, B. & Wadensjö, E. (2014), *Arbetsmarknaden*, 4 edn, Studentlittertur, Stockholm.

Borjas, G. J. (2011), 'Poverty and program participation among immigrant children', *The Future of Children* **21**(1), 247–266.

Bratsberg, B., Raaum, O. & Røed, K. (2011), 'Educating children of immigrants: Closing the gap in Norwegian schools', *Nordic Economic Policy Review* **3**(1), 211–251.

Burgess, S. & Umaña-Aponte, M. (2011), Raising your sights: the impact of friendship networks on educational aspirations, Working paper, Department of Economics, University of Bristol, UK.

Calvó-Armengol, A., Patacchini, E. & Zenou, Y. (2009), 'Peer effects and social networks in education', *The Review of Economic Studies* **76**(4), 1239–1267.

Card, D. (2005), Is the new immigration really so bad?, Working Paper 11547, National Bureau of Economic Research.

Card, D. E. & Schmidt, C. M. (2003), 'Editorial: Symposium on "second-generation immigrants and the transition to ethnic minorities"', *Journal of Population Economics: Journal of the European Society for Population Economics (ESPE)* .

Carlana, M., La Ferrara, E. & Pinotti, I. P. (2015), Shaping educational careers of immigrant children: Motivation, cognitive skills and teachers beliefs, Working paper.

Carlsson, M. & Rooth, D.-O. (2007), 'Evidence of ethnic discrimination in the Swedish labor market using experimental data', *Labour Economics* **14**(4), 716–729.

Dalton, P. S., Ghosal, S. & Mani, A. (2016), 'Poverty and aspirations failure', *The Economic Journal* **126**(590), 165–188.

Durlauf, S. (2006), 'Groups, social influences and inequality', *Poverty traps* pp. 141–175.

Engzell, P. (2016), Second generation success: What do positively selected, low-status immigrants transmit to their children?, Working paper.

Feliciano, C. & Rumbaut, R. G. (2005), 'Gendered paths: Educational and occupational expectations and outcomes among adult children of immigrants', *Ethnic and Racial Studies* **28**(6), 1087–1118.

Fordham, S. & Ogbu, J. (1986), 'Black students' school success: Coping with the burden of acting white', *The Urban Review* **18**(3), 176–206.

Fryer Jr, R. G. (2016), The production of human capital in developed countries: Evidence from 196 randomized field experiments, Technical report, National Bureau of Economic Research.

Genicot, G. & Ray, D. (2009), Aspirations, inequality, investment and mobility, Technical report, New York University.

Genicot, G. & Ray, D. (2017), 'Aspirations and inequality', *Econometrica* **85**(2), 489–519.

Goux, D., Gurgand, M. & Maurin, E. (forthcoming), 'Adjusting your dreams? The effect of school and peers on dropout behaviour', *The Economic Journal* .

Guyon, N. & Huillery, E. (2016), Aspirations failure and the escalation of social inequalities: Evidence from french teenagers, Working paper.

Halleröd, B. (2011), 'What do children know about their futures: Do children's expectations predict outcomes in middle age?', *Social Forces* **90**(1), 65–83.

Hanson, S. L. (1994), 'Lost talent: Unrealized educational aspirations and expectations among us youths', *Sociology of education* pp. 159–183.

Heath, A. & Brinbaum, Y. (2007), Guest editorial, *in* A. Heath & Y. Brinbaum, eds, 'Explaining ethnic inequalities in educational attainment', Sage Publications Sage UK: London, England, pp. 291–304.

Heath, A. & Brinbaum, Y. (2014), *Unequal Attainments: Ethnic educational inequalities in ten Western countries*, Oxford University Press.

Ichou, M. (2014), 'Who they were there: Immigrants educational selectivity and their childrens educational attainment', *European Sociological Review* **30**(6), 750–765.

Jacob, B. A. & Wilder, T. (2010), Educational expectations and attainment, Working Paper 15683, National Bureau of Economic Research.

Jonsson, J. O. & Rudolphi, F. (2011), 'Weak performancestrong determination: school achievement and educational choice among children of immigrants in Sweden', *European Sociological Review* **27**(4), 487–508.

Jonsson, J. O. & Östberg, V. (2009), 'Studying young peoples level of living: The Swedish Child-LNU', *Child Indicators Research* **3**(1), 47–64.

Joona, P. A., Gupta, N. D. & Wadensjö, E. (2014), 'Overeducation among immigrants in Sweden: incidence, wage effects and state dependence', *IZA Journal of Migration* **3**(1), 9.

Kalter, F., Heath, A. F., Hewstone, M., Jonsson, J., Kalmijn, M., Kogan, I., Van Tubergen, F., Kroneberg, C., Andersson Rydell, L., Brolin Låftman, S. et al. (2013), 'Children of immigrants longitudinal survey in four European countries (cils4eu)[dataset]'.

Kao, G. & Tienda, M. (1995), Optimism and achievement: The educational performance of immigrant youth, University of Chicago - Population Research Center, Chicago - Population Research Center.

Kruse, H. & Konstanze, J. (2016), Children of immigrants longitudinal survey in four European countries. Sociometric fieldwork report. Wave 1 2010/2011, v1.2.0. Mannheim, Technical report, Mannheim University.

Mickelson, R. A. (1990), 'The attitude-achievement paradox among black adolescents', *Sociology of Education* **63**(1), 44–61.

Mincer, J. A. (1974), *Introduction to "Schooling, Experience, and Earnings"*, Columbia University Press, pp. 1–4.

Mood, C. (2010), 'Logistic regression: Why we cannot do what we think we can do, and what we can do about it', *European Sociological Review* **26**(1), 67–82.

Mood, C. & Jonsson, J. O. (2013), Ekonomisk utsatthet och välfärd bland barn och deras familjer 19682010, Underlagsrapport till barns och ungas hälsa, vård och omsorg 2013, Socialstyrelsen.

Mora, T. & Oreopoulos, P. (2011), 'Peer effects on high school aspirations: Evidence from a sample of close and not-so-close friends', *Economics of Education Review* **30**(4), 575–581.

Morgan, S. (2006), Expectations and aspirations, *in* G. Ritzer, ed., 'Blackwell Encyclopedia of Sociology', Wiley-Blackwell Online.

Morgan, S. L. (1998), 'Adolescent educational expectations: Rationalized, fantasized, or both?', *Rationality and Society* **10**(2), 131–162.

Morgan, S. L. (2005), *On the Edge of Commitment: Educational Attainment and Race in the United States*, Stanford University Press.

Olsson, E. (2009), 'The role of relations: Do disadvantaged adolescents benefit more from high-quality social relations?', *Acta Sociologica* **52**(3), 263286.

Parrillo, V. N., ed. (2008), *Encyclopedia of Social Problems*, Sage Publications.

Portes, A. & MacLeod, D. (1999), 'Educating the second generation: Determinants of academic achievement among children of immigrants in the United States', *Journal of Ethnic and Migration Studies* **25**(3), 373–396.

Portes, A. & Rumbaut, R. G. (2001), *Legacies: The Story of the Immigrant Second Generation*, University of California Press.

Raven, J. J. (2003), Raven progressive matrices, *in* 'Handbook of nonverbal assessment', Springer, pp. 223–237.

Ray, D. (2006), Aspirations, poverty and economic change, *in* A. V. Banerjee, R. Bénabou & Mookherjee, eds, 'Understanding Poverty', Oxford Scholarship Online.

Roth, T. & Salikutluk, Z. (2012), 'Attitudes and expectations: do attitudes towards education mediate the relationship between social networks and parental expectations?', *British Journal of Sociology of Education* **33**(5), 701–722.

Rudolphi, F. (2014), *Tro, hopp och utbildning. Sociala skillnader i ungdomars utbildningsaspirationer*, Liber AB, chapter 5, pp. 125–155.

Salikutluk, Z. (2016), 'Why do immigrant students aim high?. Explaining the aspiration-achievement paradox of immigrants in germany', *European Sociological Review* .

SCB (2012), Teknisk rapport en beskrivning av genomförande och metoder, Technical report.

Schnepf, S. V. (2007), 'Immigrants educational disadvantage: an examination across ten countries and three surveys', *Journal of Population Economics* **20**(3), 527–545.

Schultz, T. W. (1960), 'Capital formation by education', *Journal of Political Economy* **68**(6), 571–583.

Szulkin, R. & Jonsson, J. O. (2007), *Ethnic segregation and educational outcomes in Swedish comprehensive schools*, Stockholm University Linnaeus Center.

Figure 1: Number of best friends in grade 8, CILS4EU
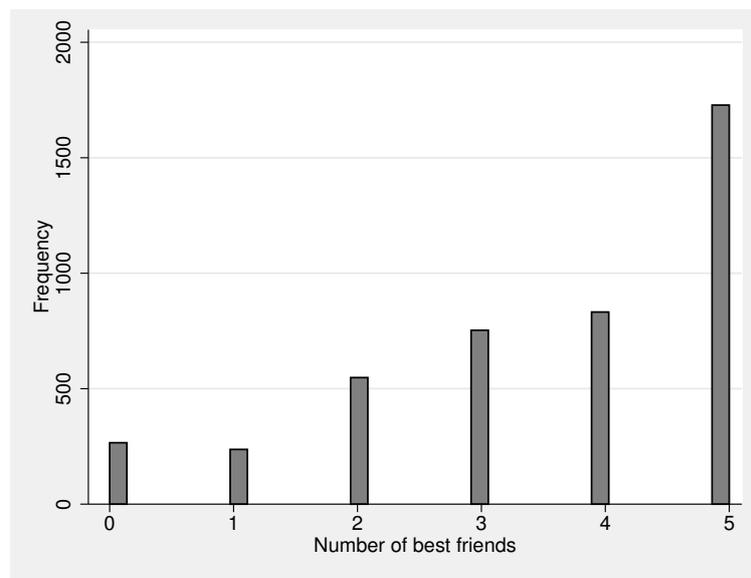
Table 1: Descriptive Statistics, LNU 2010 and CILS4EU (wave 1 and 2)

Variable	Mean	Std. Dev.	Min.	Max.	N
PANEL A: LNU 2010					
Demographics					
Age	15.919	1.678	12	19	874
Female	0.509	0.5	0	1	874
Grade	9.041	1.16	7	12	874
Parental eduction					
Parents' highest educ=Compulsory school or less	0.054	0.226	0	1	870
Parents' highest educ=High-school or less	0.286	0.452	0	1	870
Parents' highest educ=University studies	0.66	0.474	0	1	870
Self-assessments					
Self-assessment=Best in class	0.035	0.183	0	1	869
Self-assessment=Among the best	0.334	0.472	0	1	869
Self-assessment=Better than the majority	0.221	0.415	0	1	869
Self-assessment=About as good as most people	0.379	0.485	0	1	869
Self-assessment=Not as good as most people	0.032	0.177	0	1	869
Family characteristics					
Intact family	0.706	0.456	0	1	874
Reconstituted family	0.142	0.349	0	1	874
Single parent family	0.146	0.354	0	1	874
Household members	4.261	1.221	2	11	874
Educational aspirations and expectations					
University aspirations	0.768	0.423	0	1	874
University expectations	0.787	0.41	0	1	874
Aspirations-expectations gap	0.026	0.16	0	1	874
PANEL B: CILS4EU					
Demographics					
Female	0.514	0.5	0	1	4364
Foreign-born parents	0.305	0.46	0	1	4364
Parental education					
Higher educated parents	0.434	0.496	0	1	4364
Academic potential					
Language test score (std)	0.078	0.954	-3.524	2.266	4346
Cognitive test score (std)	0.055	0.971	-3.526	1.911	4334
School achievement					
Incomplete grades: Maths	0.068	0.251	0	1	3601
Incomplete grades: Swedish	0.035	0.185	0	1	3595
Incomplete grades: English	0.045	0.207	0	1	4364
Educational aspirations and expectations					
University aspirations	0.667	0.471	0	1	4364
University expectations	0.522	0.5	0	1	4364
Aspirations-expectations gap	0.177	0.382	0	1	4364
Parent's aspirations and expectations					
Parent consistency	0.307	0.461	0	1	4364
Parent discrepancy	0.225	0.418	0	1	4364
Friendship network characteristics					
Prop. female friends	0.51	0.337	0	1	4091
Prop. friends with foreign background	0.31	0.371	0	1	4091
Prop. friends with higher educ parents	0.425	0.331	0	1	4091
Prop. university aspirations friends	0.651	0.317	0	1	4091

Figure 2: Characteristics of best friends in grade 8, CILS4EU

(a)
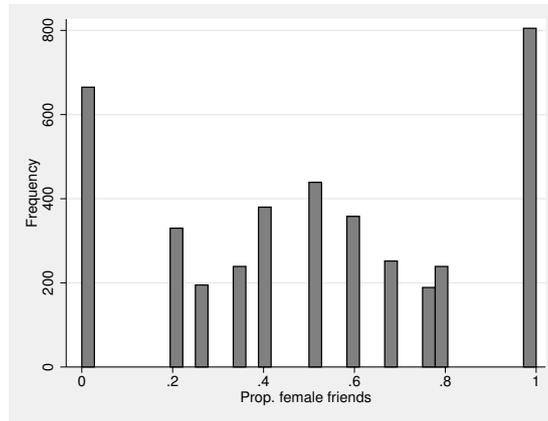
(b)
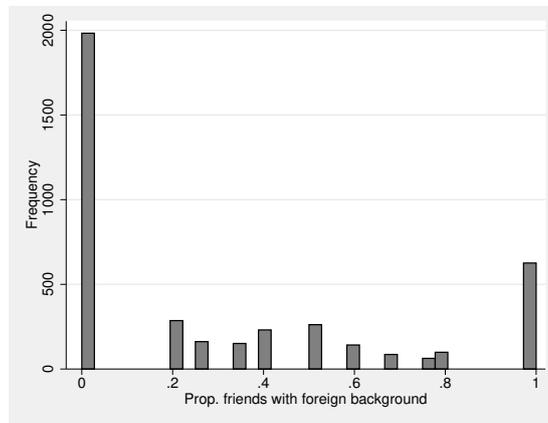
(c)
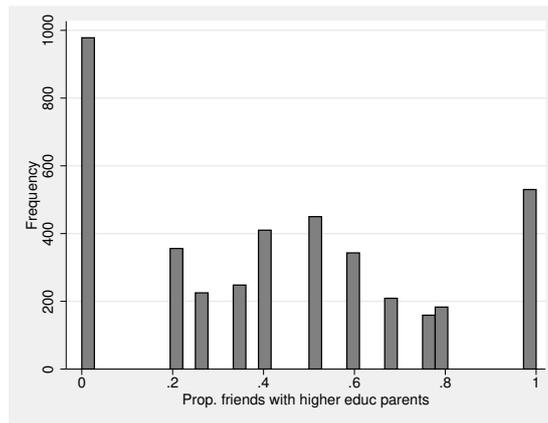
Figure 3: Destination language test score distributions in grade 8 by parents' immigration status, CILS4EU
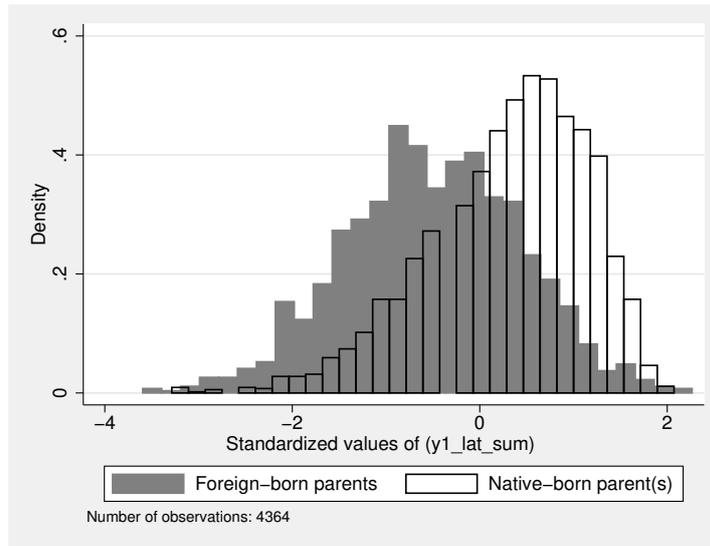
Figure 4: Cognitive ability test score distributions in grade 8 by parents' immigration status, CILS4EU
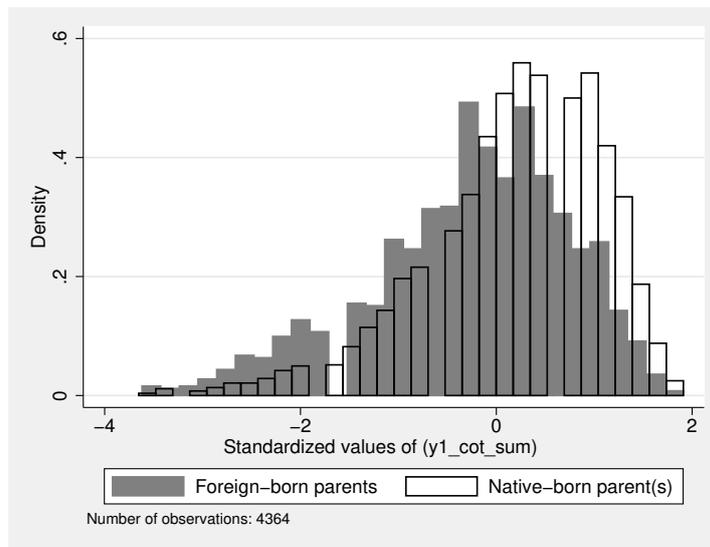
Figure 5: Aspirations by parents' immigration status, LNU 2010

Educational aspirations

Native-born parent(s)
- No, absolutely not — 10.3
- No, probably not — 20.9
- Yes, probably — 31.8
- Yes, absolutely — 37.0

Foreign-born parents
- No, absolutely not — 2.9
- No, probably not — 8.5
- Yes, probably — 26.3
- Yes, absolutely — 62.4

Number of observations: 894
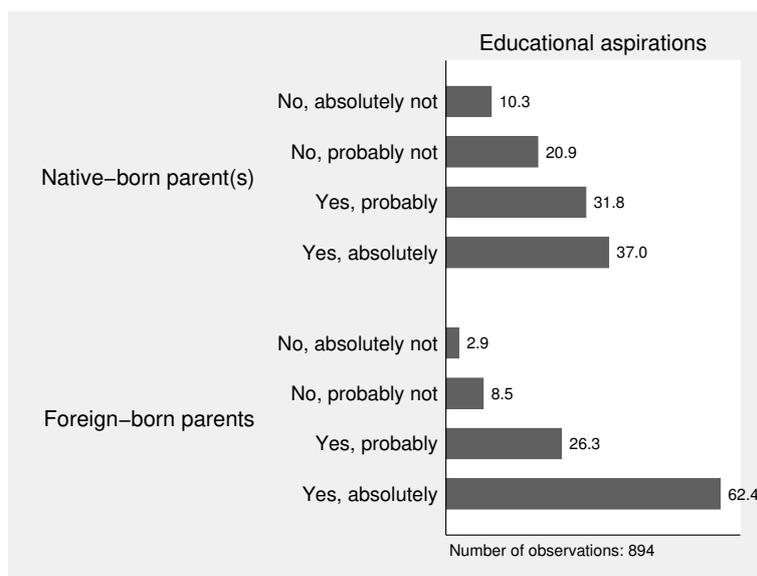
Figure 6: Expectations by parents' immigration status, LNU 2010

Educational expectations

Native-born parent(s)
- No, absolutely not — 8.7
- No, probably not — 19.7
- Yes, probably — 37.3
- Yes, absolutely — 34.3

Foreign-born parents
- No, absolutely not — 1.1
- No, probably not — 9.9
- Yes, probably — 37.7
- Yes, absolutely — 51.4

Number of observations: 894

43
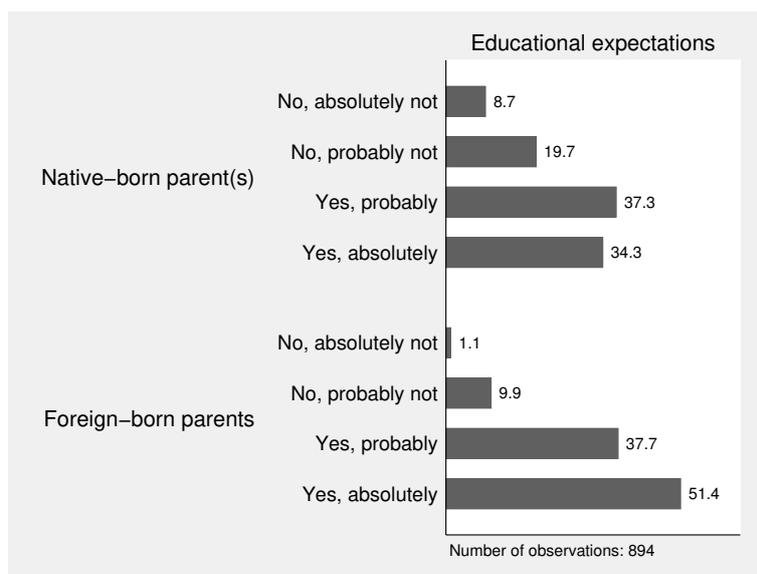
Table 2: Expectations by aspirations (%) among children with native-born parents, frequencies (weighted) in parenthesis

	Aspirations				
Expectations	No, absolutely not	No, probably not	Yes, probably	Yes, absolutely	Total
No, absolutely not	**71**	5	0	1	9
	(40.47)	(6.34)	(0)	(1.33)	(48.14)
No, probably not	24	**72**	7	0	20
	(13.75)	(83.42)	(12.50)	(0)	(109.67)
Yes, probably	5	23	**76**	21	37
	(3.01)	(26.36)	(134.95)	(42.96)	(207.28)
Yes, absolutely	0	0	17	**78**	34
	(0)	(0)	(29.40)	(161.50)	(190.90)
Total	100	100	100	100	100
	(57.23)	(116.12)	(176.86)	(205.79)	(556)

Note: In order to adjust for family size we use the weights provided in the technical report by SCB (see details in SCB (2012)).

Table 3: Expectations by aspirations (%) among children with foreign-born parents, frequencies (weighted) in parenthesis

	Aspirations				
Expectations	No, absolutely not	No, probably not	Yes, probably	Yes, absolutely	Total
No, absolutely not	**37**	0	0	0	1
	(3.38)	(0)	(0)	(0)	(3.38)
No, probably not	6	**74**	9	2	10
	(.60)	(19.84)	(7.19)	(3.80)	(31.43)
Yes, probably	57	16	**75**	24	38
	(5.26)	(4.29)	(62.50)	(47.79)	(119.85)
Yes, absolutely	0	10	17	**74**	51
	(0)	(2.79)	(13.81)	(146.74)	(163.34)
Total	100	100	100	100	100
	(9.25)	(26.93)	(83.50)	(198.32)	(318)

Note: In order to adjust for family size we use the weights provided in the technical report by SCB (see details in SCB (2012)).

Figure 7: Aspirations grade 8 by parents' immigration status, CILS4EU
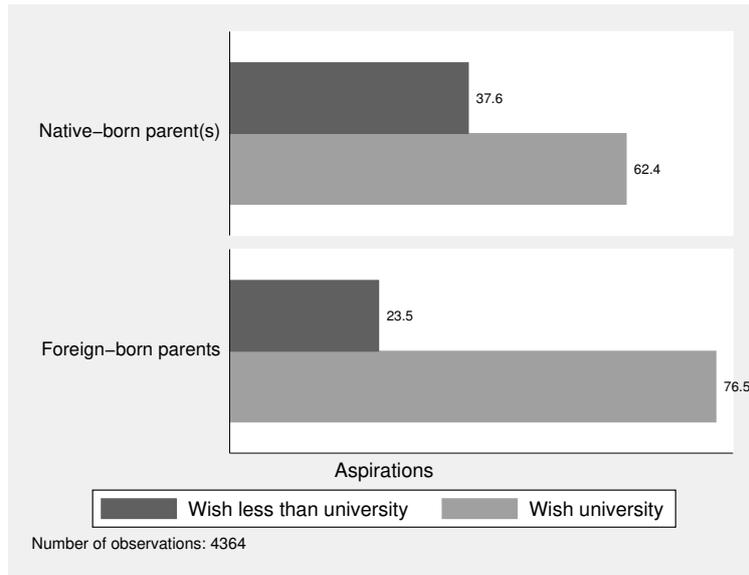
Figure 8: Expectations grade 8 by parents' immigration status, CILS4EU
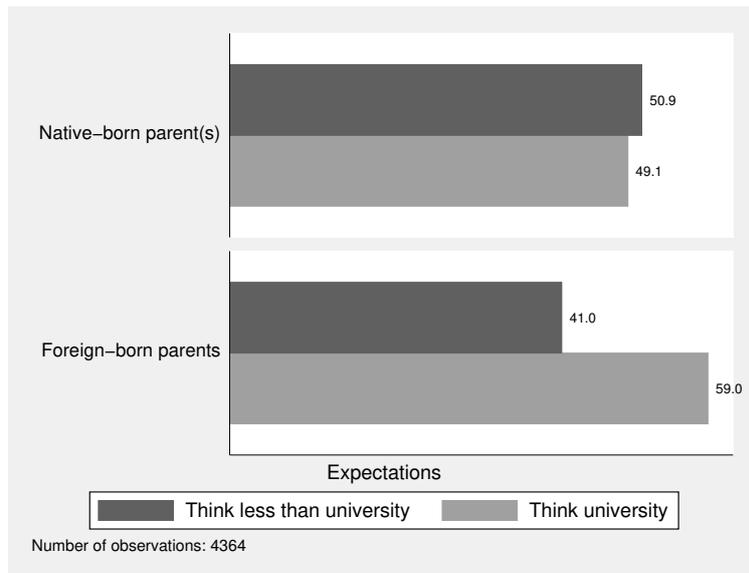
Figure 9: Aspirations-expectations gap grade 8 by parents' immigration status, CILS4EU
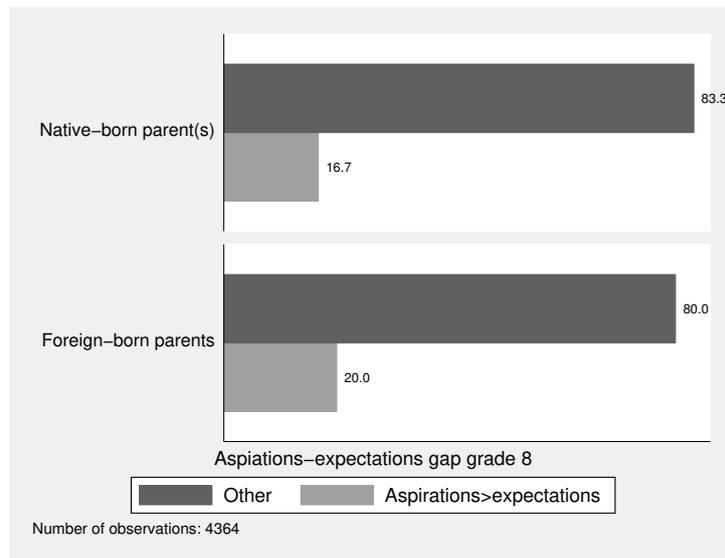
Figure 10: Aspirations-expectations gap grade 9 by parents' immigration status, CILS4EU
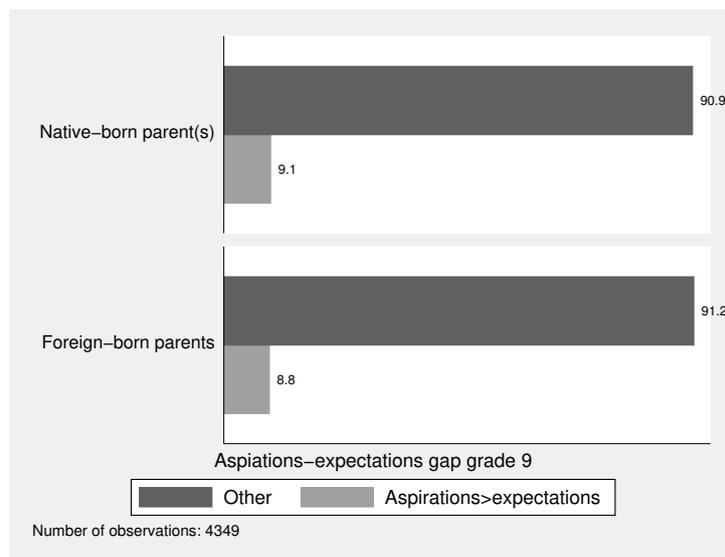
Table 4: OLS coefficients from models of aspirations, expectations and the aspirations-expectations gap, LNU

	Aspirations			Expectations			Aspirations-expectations gap		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Foreign-born parents	0.178***	0.255***		0.157***	0.227***		0.00312	0.0430	
	(0.029)	(0.054)		(0.028)	(0.054)		(0.011)	(0.034)	
Age (demeaned)		−0.00682	−0.00661		−0.0129	−0.0125		−0.00126	−0.00105
		(0.009)	(0.009)		(0.009)	(0.009)		(0.004)	(0.004)
Female		0.216***	0.185***		0.217***	0.189***		−0.0178	−0.0259**
		(0.040)	(0.034)		(0.038)	(0.033)		(0.013)	(0.013)
Foreign-born parents × Female		−0.171***			−0.157**			−0.0428	
		(0.062)			(0.065)			(0.036)	
Parents' highest educ=University studies		0.180***	0.188***		0.191***	0.198***		−0.00250	−0.00174
		(0.037)	(0.038)		(0.036)	(0.036)		(0.013)	(0.013)
Self-assessment=Best in class		0.396***	0.369***		0.356***	0.331***		−0.0259	−0.0318
		(0.118)	(0.117)		(0.113)	(0.112)		(0.037)	(0.038)
Self-assessment=Among the best		0.353***	0.350***		0.285***	0.281***		−0.00669	−0.00833
		(0.094)	(0.094)		(0.096)	(0.096)		(0.037)	(0.037)
Self-assessment=Better than the majority		0.375***	0.375***		0.314***	0.314***		−0.0256	−0.0259
		(0.095)	(0.094)		(0.095)	(0.095)		(0.038)	(0.038)
Self-assessment=About as good as most people		0.227**	0.229**		0.178*	0.180*		−0.00295	−0.00337
		(0.095)	(0.094)		(0.094)	(0.094)		(0.038)	(0.038)
Reconstituted family		0.0248	0.0326		0.00904	0.0158		0.0218	0.0229
		(0.046)	(0.047)		(0.045)	(0.045)		(0.019)	(0.019)
Single parent family		−0.0174	−0.0138		−0.0101	−0.00448		−0.0170	−0.0152
		(0.052)	(0.052)		(0.051)	(0.051)		(0.016)	(0.015)
Household members		0.0149	0.0103		0.00872	0.00570		−0.00739	−0.00738*
		(0.014)	(0.014)		(0.014)	(0.015)		(0.005)	(0.004)
European			0.0841			0.0755*			0.0143
			(0.053)			(0.043)			(0.026)
non-European			0.214***			0.186***			0.0242
			(0.049)			(0.047)			(0.021)
Constant	0.703***	0.115	0.144	0.730***	0.219*	0.240**	0.0252***	0.0741	0.0781
	(0.018)	(0.111)	(0.110)	(0.017)	(0.114)	(0.114)	(0.007)	(0.048)	(0.048)
Observations	876	867	867	874	865	865	874	865	865
Adjusted R^2	0.040	0.149	0.146	0.033	0.145	0.142	−0.001	0.004	0.001

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: OLS regressions with robust standard errors clustered at the family level. In Models (1)-(3) the dependent variable is a dummy defined as having aspirations for a university degree (i.e. having responded "Yes, absolutely" or "Yes, probably" on the question "Would you like to continue going to school after the upper secondary level, that is, attend a university or university college?". The dependent variable in Models (4)-(6) is a dummy defined as having expectations for a university degree (i.e. having responded "Yes, absolutely" or "Yes, probably" on the question "Do you think you will actually continue going to school after the upper secondary level, that is, attend a university or university college?"). In Models (7)-(9) the dependent variable is a dummy for the aspirations-expectations gap which is defined as having aspirations for a university degree and but not expecting to get one (i.e. having responded "Yes, absolutely" or "Yes, probably" on the aspirations question and "No, probably not" or "No, absolutely not" on the expectations question.

Figure 11: Grade distributions in core subjects by parents' immigration status, self-reported in grade 9, CILS4EU

(a)

Grade Maths

Number of observations: 4120

(b)

Grade Swedish

Number of observations: 4237

(c)

Grade English

Number of observations: 4168
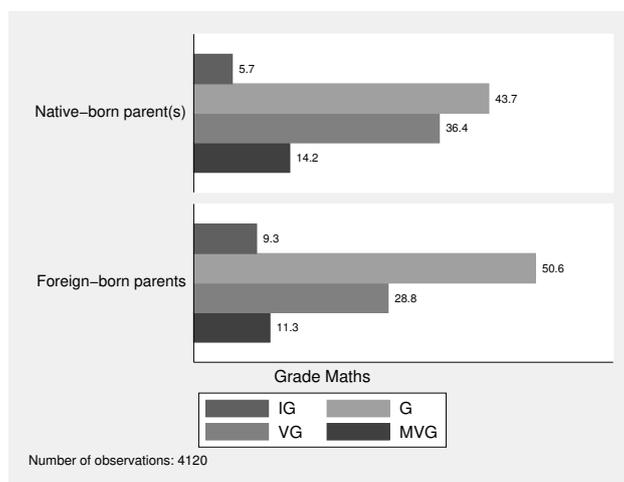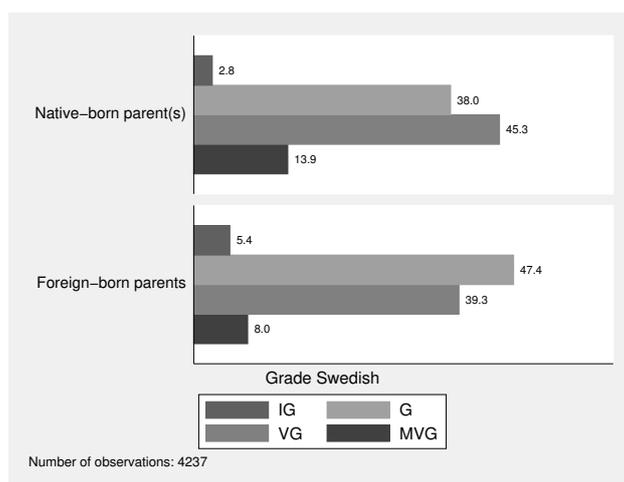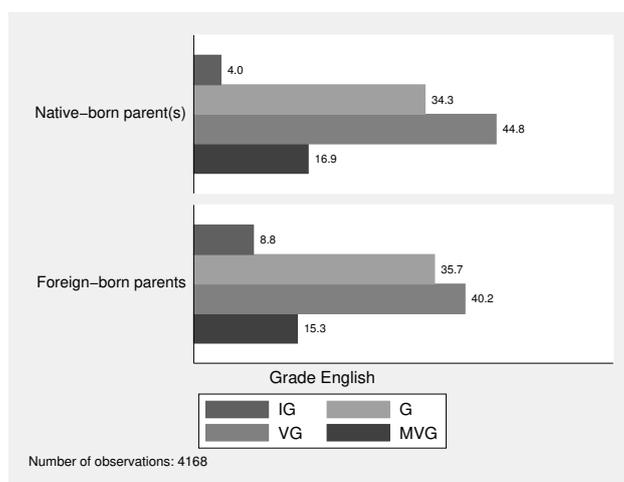
Table 5: OLS coefficients from regressions of aspirations in grade 8 and 9 respectively, CILS4EU

	Aspirations grade 8			Aspirations grade 9			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Foreign-born parents	0.107***	0.117***	0.257***	0.0946***	0.133***	0.235***	0.234***
	(0.021)	(0.022)	(0.028)	(0.021)	(0.021)	(0.027)	(0.027)
Female			0.108***			0.120***	0.115***
			(0.021)			(0.019)	(0.020)
Foreign-born parents × Female			−0.0756**			−0.0796**	−0.0814***
			(0.033)			(0.031)	(0.031)
Higher educated parents			0.105***			0.0644***	0.0641***
			(0.016)			(0.014)	(0.014)
Language test score (std)			0.116***			0.0310***	0.0319***
			(0.010)			(0.010)	(0.010)
Cognitive test score (std)			0.0400***			0.0146	0.0144
			(0.010)			(0.010)	(0.010)
All friends: female			−0.0107				0.0253
			(0.025)				(0.020)
All friends: foreign-born parents			0.0441				0.0466
			(0.034)				(0.032)
All friends: higher educ parents			0.0152				−0.00678
			(0.026)				(0.025)
All friends: university aspirations			−0.0390*				−0.0228
			(0.021)				(0.019)
Grade sum						0.0721***	0.0723***
						(0.005)	(0.005)
Classroom FEs	✓	✓	✓	✓	✓	✓	✓
Observations	4364	4075	4075	4354	3381	3381	3381
Adjusted R^2	0.007	0.009	0.102	0.005	0.013	0.184	0.184

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: OLS regressions. Standard errors are clustered at the classroom level. All regressions include classroom fixed-effects. Variable definitions are found in section 3.7. The dependent variable is a dummy defined as having aspirations for a university degree ("What is the highest level of education you wish to get? (Don't know, No degree, Compulsory school, Upper secondary school, College/university)").

Table 6: OLS coefficients from regressions of expectations in grade 8 and 9 respectively, CILS4EU

	Expectations grade 8			Expectations grade 9			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Foreign-born parents	0.0714***	0.0802***	0.224***	0.0981***	0.135***	0.246***	0.245***
	(0.021)	(0.023)	(0.028)	(0.022)	(0.023)	(0.028)	(0.028)
Female			0.0710***			0.119***	0.114***
			(0.021)			(0.019)	(0.019)
Foreign-born parents × Female			−0.0667*			−0.0793**	−0.0806**
			(0.035)			(0.034)	(0.033)
Higher educated parents			0.102***			0.106***	0.105***
			(0.017)			(0.015)	(0.015)
Language test score (std)			0.122***			0.0292***	0.0288***
			(0.010)			(0.010)	(0.010)
Cognitive test score (std)			0.0593***			0.0135	0.0137
			(0.009)			(0.009)	(0.009)
All friends: female			0.0277				0.0239
			(0.028)				(0.021)
All friends: foreign background			0.0381				0.0231
			(0.040)				(0.041)
All friends: higher educ parents			−0.0231				−0.0522*
			(0.029)				(0.028)
All friends: university aspirations			−0.0110				0.0135
			(0.024)				(0.020)
Grade sum						0.0899***	0.0902***
						(0.005)	(0.005)
Classroom FEs	✓	✓	✓	✓	✓	✓	✓
Observations	4364	4075	4075	4353	3377	3377	3377
Adjusted R^2	0.003	0.003	0.098	0.005	0.011	0.224	0.225

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: OLS regressions. Standard errors are clustered at the classroom level. All regressions include classroom fixed-effects. Variable definitions are found in section 3.7. The dependent variable is a dummy defined as having expectations for a university degree ("What is the highest level of education think you will actually get? (Don't know, No degree, Compulsory school, Upper secondary school, College/university)").

Table 7: OLS coefficients from regressions predicting the aspirations-expectations gap in grade 8 and 9 respectively, CILS4EU

	Aspirations-expectations gap grade 8			Aspirations-expectations gap grade 9			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Foreign-born parents	0.0312**	0.0316*	0.0242	−0.00503	−0.00629	−0.00934	−0.00846
	(0.015)	(0.016)	(0.021)	(0.012)	(0.015)	(0.019)	(0.019)
Female			0.0272*			0.0102	0.0103
			(0.016)			(0.013)	(0.013)
Foreign-born parents × Female			−0.00864			−0.0207	−0.0210
			(0.027)			(0.025)	(0.025)
Higher educated parents			−0.00763			−0.0407***	−0.0402***
			(0.013)			(0.012)	(0.012)
Language test score (std)			−0.00928			−0.00215	−0.00120
			(0.008)			(0.008)	(0.008)
Cognitive test score (std)			−0.0146**			−0.000467	−0.000799
			(0.007)			(0.007)	(0.007)
All friends: female			−0.0322				0.00229
			(0.021)				(0.017)
All friends: foreign-born parents			0.0180				0.0178
			(0.035)				(0.029)
All friends: higher educ parents			0.0140				0.0343*
			(0.021)				(0.021)
All friends: university aspirations			−0.0200				−0.0260*
			(0.020)				(0.014)
Grade sum						−0.0187***	−0.0188***
						(0.003)	(0.003)
Classroom FEs	✓	✓	✓	✓	✓	✓	✓
Observations	4364	4075	4075	4349	3377	3377	3377
Adjusted R^2	0.001	0.001	0.003	−0.000	−0.000	0.018	0.019

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: OLS regressions. Standard errors are clustered at the classroom level. All models include classroom fixed-effects. Variable definitions are found in section 3.7. The dependent variable is a dummy defined as having aspirations for a university degree (i.e. having responded "College/university" on the question "What is the highest level of education you wish to get? (Don't know, No degree, Compulsory school, Upper secondary school, College/university)" but not expecting to get one (i.e. having responded less than "College/university" on the question "What is the highest level of education you think you will actually get? (Don't know, No degree, Compulsory school, Upper secondary school, College/university)" .

Table 8: OLS coefficients from models of aspirations, expectations and the aspirations-expectations gap in grade 8 and 9 including parent discrepancy, CILS4EU

|	Aspirations		Expectations		Aspirations-expectations gap	
	Grade 8	Grade 9	Grade 8	Grade 9	Grade 8	Grade 9
Foreign-born parents	0.226***	0.176***	0.201***	0.244***	0.0202	−0.0567**
	(0.038)	(0.033)	(0.042)	(0.035)	(0.032)	(0.024)
Female	0.0968***	0.106***	0.0679**	0.136***	0.0217	−0.00916
	(0.026)	(0.022)	(0.027)	(0.024)	(0.020)	(0.016)
Foreign-born parents × Female	−0.0665	−0.0534	−0.0607	−0.119***	−0.00805	0.0323
	(0.045)	(0.038)	(0.050)	(0.044)	(0.040)	(0.034)
Higher educated parents	0.110***	0.0568***	0.119***	0.0937***	−0.0212	−0.0344**
	(0.019)	(0.019)	(0.022)	(0.020)	(0.018)	(0.016)
Language test score (std)	0.124***	0.0142	0.117***	0.0107	−0.00343	−0.00425
	(0.014)	(0.013)	(0.014)	(0.014)	(0.012)	(0.011)
Cognitive test score (std)	0.0205*	0.000393	0.0433***	0.00929	−0.0166	−0.00934
	(0.012)	(0.013)	(0.012)	(0.013)	(0.011)	(0.010)
Parent discrepancy	−0.0400**	−0.0436**	−0.116***	−0.0714***	0.0629***	0.0343**
	(0.019)	(0.018)	(0.020)	(0.020)	(0.018)	(0.016)
All friends: female	−0.0192	−0.00957	0.0219	−0.00440	−0.0389	−0.00312
	(0.032)	(0.025)	(0.036)	(0.028)	(0.025)	(0.023)
All friends: foreign-born parents	0.0989*	0.0427	0.0328	0.0326	0.0731*	−0.00102
	(0.051)	(0.050)	(0.052)	(0.057)	(0.044)	(0.038)
All friends: higher educ parents	−0.00147	−0.0323	−0.0263	−0.0834**	0.00419	0.0409
	(0.035)	(0.028)	(0.038)	(0.034)	(0.027)	(0.025)
All friends: university aspirations	−0.0520*	−0.0407*	0.00366	0.0124	−0.0540**	−0.0458**
	(0.028)	(0.024)	(0.030)	(0.025)	(0.024)	(0.018)
Grade sum		0.0709***		0.0880***		−0.0187***
		(0.006)		(0.006)		(0.005)
Classroom FEs	✓	✓	✓	✓	✓	✓
Observations	2438	2139	2438	2135	2438	2135
Adjusted R^2	0.097	0.166	0.106	0.227	0.012	0.030

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: OLS regressions. Standard errors are clustered at the classroom level. All models include classroom fixed-effects. Variable definitions are found in section 3.7.

Figure 12: Complete/incomplete grades in grade 9 by immigrant status, CILS4EU

Native–born parent(s)
91.1
8.9

Foreign–born parents
84.3
15.7

Complete grades   Incomplete grades

Number of observations: 3587
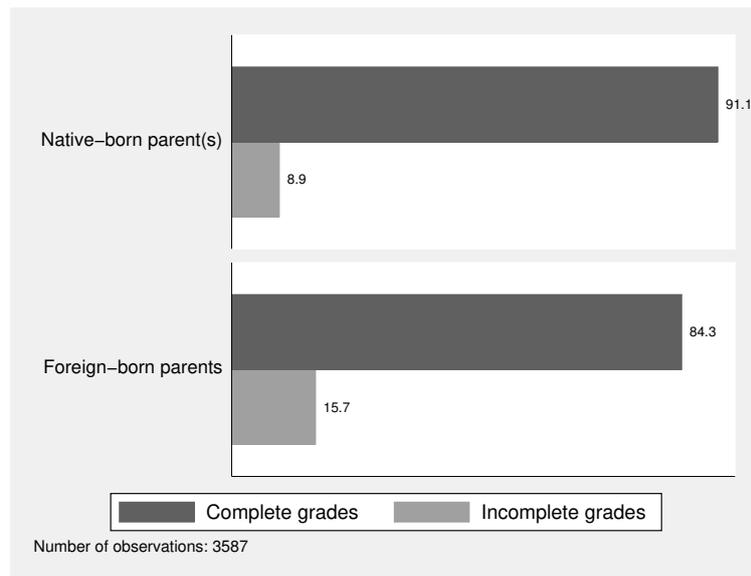
Table 9: Subgroup analysis of the aspirations-expectations gap in grade 9 among individuals with complete and incomplete grades, OLS coefficients and standard errors in parenthesis

	Aspirations-expectations gap		
	(1)	(2)	(3)
Panel A: Non-complete grades			
Foreign-born parents	0.0201	0.00993	−0.0206
	(0.042)	(0.120)	(0.115)
Female		0.0244	0.0170
		(0.108)	(0.108)
Foreign-born parents × Female		−0.0205	−0.0293
		(0.128)	(0.135)
Higher educated parents		−0.0425	−0.0291
		(0.082)	(0.084)
Language test score (std)		0.0317	0.0410
		(0.031)	(0.031)
Classroom FEs	✓	✓	✓
Observations	351	351	351
Adjusted R^2		−0.007	0.006
Panel B: At least Pass in all core subjects			
Foreign-born parents	−0.00834	−0.0300	−0.0284
	(0.012)	(0.020)	(0.020)
Female		−0.00297	−0.00345
		(0.012)	(0.013)
Foreign-born parents × Female		−0.0116	−0.0110
		(0.025)	(0.025)
Higher educated parents		−0.0421***	−0.0417***
		(0.012)	(0.012)
Language test score (std)		−0.0231***	−0.0223***
		(0.008)	(0.008)
Classroom FEs	✓	✓	✓
Observations	3034	3034	3034
Adjusted R^2		0.008	0.009

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Linear probability models (LPM) with classroom fixed-effects and standard errors clustered at the classroom level. While column (1) displays the "raw" immigrant-non-immigrant difference, column (2) includes a set of individual-level explanatory variables. The model in column (3) also controls for friends' average characteristics including foreign background, female, higher educated parents and aspirations as defined in section 3.7.

Table 10: Predictors of lost talent, alternative definitions, CILS4EU

	Grade 8				Grade 9			
	Cutoff A		Cutoff B		Cutoff A		Cutoff B	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Foreign-born parents	−0.0170*	0.00394	−0.0189**	−0.00979	−0.0142	0.00298	−0.0142	−0.00293
	(0.010)	(0.015)	(0.008)	(0.012)	(0.009)	(0.012)	(0.009)	(0.009)
Female		0.0261**		0.0111		0.0144		0.000995
		(0.013)		(0.010)		(0.009)		(0.008)
Female × Foreign-born parents		−0.0139		0.00806		−0.0220		−0.00653
		(0.019)		(0.016)		(0.015)		(0.013)
Higher educated parents		−0.00941		−0.0102		−0.0126		−0.00355
		(0.010)		(0.009)		(0.009)		(0.007)
Language test score (std)		0.0285***		0.0276***		0.0102*		0.0106**
		(0.005)		(0.004)		(0.005)		(0.004)
All friends: female		−0.0241*		−0.0249**		−0.000246		−0.00377
		(0.013)		(0.010)		(0.010)		(0.007)
All friends: foreign-born parents		0.0184		0.0204*		−0.00348		0.00201
		(0.015)		(0.012)		(0.013)		(0.010)
All friends: higher educ parents		−0.0196		0.00334		0.0145		0.0112
		(0.014)		(0.013)		(0.012)		(0.010)
All friends: university aspirations		0.00200		0.00161		−0.00199		−0.0134**
		(0.011)		(0.010)		(0.008)		(0.007)
Grade sum						−0.00552**		−0.00211
						(0.002)		(0.002)
Constant	0.105***	0.0922***	0.0711***	0.0637***	0.0618***	0.0836***	0.0618***	0.0523***
	(0.006)	(0.009)	(0.005)	(0.008)	(0.005)	(0.013)	(0.005)	(0.011)
Observations	4075	4075	4075	4075	3377	3377	3377	3377
Adjusted R^2	0.000	0.008	0.001	0.010	0.000	0.002	0.000	0.001

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Linear probability model (LPM) with standard errors are clustered at the classroom level. Variable definitions are found in section 3.7. The dependent variable is a dummy for lost talent which is defined as scoring higher than the mean or the median of the sample on the cognitive ability test administered in grade 8 and expressing an aspirations-expectations gap. In columns (1)-(2) and (5)-(6) the threshold is defined as a cognitive ability test score higher than the mean in the analysis sample. In columns (3)-(4) and (7)-(8) the threshold is defined as a test score higher than the median in the analysis sample.

Appendices

A Model specification

In this section, I check if the results are robust to changing model specifications. My preferred model is the OLS fixed-effects model. As a robustness test I also estimate an OLS random-effects model. The random or fixed intercepts handle the clustered nature of my data. I perform a Hausman test to see whether the fixed and random-effects are significantly different. The test result shows that I can reject the null, i.e. a fixed-effects model is more suitable for the data.

The fixed-effects model is preferable to the random-effects model since the purpose of this paper is to account for differences in aspirations, expectations and the aspirations-expectations gap, and not to estimate variance components. A fixed-effects estimation allows me to disentangle the influence of classroom-specific factors from the influence of individual factors and reduces the omitted variable bias. Put differently, the fixed-effects model takes account of the correlation among individuals who belong to the same cluster, in this case the classroom.

As the results may be sensitive to the functional form of the model I present the results from separate logit and probit regressions. The latent response model corresponding to equation (1) is:

$$
\begin{aligned}
y_i^* = {} & \alpha + \beta_1 immstat_i + \beta_2 age_i + \beta_3 female_i + \beta_4 immstat_i \times female_i \\
& + \sum_{j=1}^{2} \beta_{5j} educ_{ji} + \sum_{j=1}^{3} \beta_{6j} selfass_{ji} + \sum_{j=1}^{3} \beta_{7j} famtype_{ji} + \beta_7 hhsize_i + \epsilon_i,
\end{aligned}
\tag{5}
$$

where we only observe $y_i = I(y_i^* > 0)$ for the latent variable y_i^*. Moreover, I estimate the corresponding unconditional and conditional logit and probit models. Since odds ratios are not comparable across nested models, I present the result in the form of average marginal effects whenever possible. Table A1 shows the results from logistic regressions on the LNU data.[42] In contrast to the marginal effects at means where all covariates are set to their average value, the average marginal effects uses all the data. For binary variables the coefficient indicates the predicted probabilities change as the variable goes from 0 to 1. With regard to continuous variables, the average marginal effect measures the instantaneous rate of change. The average marginal effect of a continuous variable such as age indicates by how many units (ppt) the probability to have aspirations for university education changes if the explanatory variable changes by one year. Average marginal effects are less useful when models include interaction terms and in those cases I instead use odds ratios. Table A2 shows the results from probit regressions without fixed-effects using the LNU data.

Due to the incidental parameters problem of binary choice models, unconditional logit and probit may be erroneous. Estimating these models with classroom dummies may produce inconsistent estimates (of both the fixed-effects estimate and the other coefficients). The conditional fixed-effect logit model deals with this problem, however, the classroom-effect is not estimated in the fixed-effect logit (xtlogit

---

[42]See Mood (2010) for a discussion on comparing odds ratios across nested models.

command in Stata). The functional form of the logit allows for elimination of the classroom-specific term in the conditional fixed-effects logit case but there is the problem of interpreting the effects: marginal effects cannot be estimated unless one assumes the constants are 0.

Another alternative to the unconditional logit and probit models is to use the probit model with random-effects where marginal effects can be estimated at constants equaling null. This model does, however, not allow for correlation between the classroom-specific effect and any of the explanatory variables. As such it does not handle the endogeneity issue. Table A3 shows the odds ratios from random-effects probit regressions.

Tables A4, A5 and A6 show the results from random-effects logistic regressions of aspirations, expectations and the gap. There is no command in Stata for conditional fixed-effects probit estimation. Tables A7, A8 and A9 show the results from fixed-effects logistic regressions.

Tables A10, A11, A12 and A13 present the results from multilevel logit regressions. It is not possible to calculate average marginal effects using the Stata command melogit due to the mixed design of the model with both random (individual effects) and fixed effects (explanatory variables). However, predicted probabilities can be calculated separately using marginsplots for specific categories of the independent variables. These results are then analyzed using odds ratios. The results in this section are presented both as log odds and odds ratios for ease of interpretation. Overall, the results from the different specifications display a similar picture and are largely confirmatory.

B   Ordered logit and probit models

As previously mentioned, dichotomizing the dependent variable may lead to a loss of important information from the model. The primary reason for dichotomizing the dependent variables is comparability between the two samples. Moreover, the character of the CILS4EU data makes it less suitable for an ordered model which is why the following sensitivity analysis is based on the LNU data. In order to see whether any interesting information is lost from dichotomizing the dependent variables in the LNU data I also perform ordered logit and probit regressions. In contrast to OLS which assumes that a movement from 1 to 2 is equivalent to 3 from good to 4 on the aspirations/expectations scale the ordered logit and probit models take account of the thresholds of the underlying continuous latent variable.

First, I estimate the two models separately. I only run the ordered logit and probit models for aspirations and expectations on the LNU data and not the CILS4EU data since the former is more suitable for these types of models. In the LNU questionnaires the respondents have been asked if they would you like to continue going to school after the upper secondary level and whether they think they will actually continue going to school after the upper secondary level with ordered response options "Yes, absolutely", "Yes, probably", "No, probably not" and "No, absolutely not". In the CILS4EU questionnaire they have been asked to specify the level of education they think or wish they will get. The results from ordered logit and probit regressions are presented in table B1. The coefficients of the variable foreign-born

parents from the two models are positive and significant (the magnitude of the coefficients cannot be compared since they are on different scales). Having foreign-born parents makes one significantly more likely to belong the upper categories of the dependent variables aspirations and expectations.

First, I assume that the underlying variable y_i^* is continuous and interval level and the error term is independently logistically distributed. The bivariate ordered probit model takes into account the multiple ordered response categories (there are 16 mutually exclusive outcomes). The bivariate ordered probit model is an extension of the bivariate probit model. The results are presented in table B2. There are too few observations in the lowest and highest categories to get meaningful results out of an ordered logit or probit model of the aspirations-expectations gap. The bivariate model is the correct specification since the "rho" parameter is statistically significant in the baseline model with an immigrant dummy (95 percent confidence interval: [0.86; 0.93]). This means that the error terms in the two equations are correlated. The sign of the estimated coefficients for the immigrant dummy are positive for both aspirations and expectations indicating that the latent variable y_i^* increases with this regressor. The lower panel of the table shows the estimates of the thresholds.

Tables B3 and B4 show the predicted joint probabilities from bivariate ordered regressions (SUR) of aspirations and expectations for children with native-born parents and children with foreign-born parents (unconditional on background variables). The table cells add up to 100 percent. As suggested by the results from the ordered logit and probit regression results, children with immigrant parent(s) are more likely to belong to the lower-right corner category of the table. An important next step is to investigate whether the results hold for different definitions of immigrant children.

C   Definitions of immigrant children

In this section, I check if the results are robust to changing variable definitions. In table C1, C2 and C3 I define immigrant children as children who were born abroad. The reference category consists of boys who are native-born and/or have native-born parents. Conditional on individual, family and network characteristics, children born abroad are more likely to have university aspirations than their peers (12.2 ppt for immigrant boys and 13.7 ppt for immigrant girls, column (3)). With regard to expectations, a similar picture emerges from table C2. Furthermore, the immigrant-non-immigrant differentials in aspirations and expectations seem to grow over time. I find no significant differences in the immigrant-non-immigrant aspirations-expectations gap presented in column (3) and (7) suggesting that immigrant children are not more likely to have expectations falling short of aspirations than their peers.

Table A1: Odds ratios and average marginal effects from logistic regressions of aspirations, expectations and the aspirations-expectations gap by region of origin, LNU

| | Educational aspirations | | | | | | Educational expectations | | | | | | Aspirations-expectations gap | | | | | |
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | | Model 6 | | Model 7 | | Model 8 | | Model 9 | |
	OR	AME	OR	AME	OR	AME	OR	AME	OR	AME	OR	AME	OR	AME	OR	AME	OR	AME
Foreign-born parents	3.132***	0.178***	4.361***	0.173***			2.894***	0.157***	3.821***	0.152***			1.128	0.00312	3.009*	0.0252		
	(0.657)	(0.028)	(1.754)	(0.040)			(0.582)	(0.026)	(1.553)	(0.038)			(0.488)	(0.011)	(1.803)	(0.021)		
Age (demeaned)			0.962	-0.00650	0.961	-0.00669			0.922	-0.0128	0.921	-0.0129			0.952	-0.00125	0.954	-0.00120
			(0.054)	(0.009)	(0.053)	(0.009)			(0.054)	(0.009)	(0.053)	(0.009)			(0.149)	(0.004)	(0.148)	(0.004)
Female			3.192***	0.184***	3.022***	0.186***			3.459***	0.188***	3.301***	0.190***			0.458	-0.0266**	0.339**	-0.0265**
			(0.706)	(0.033)	(0.635)	(0.034)			(0.798)	(0.032)	(0.719)	(0.033)			(0.256)	(0.013)	(0.172)	(0.013)
Foreign-born parents × Female			0.515						0.561						0.238			
			(0.282)						(0.339)						(0.295)			
Parents' highest educ=University studies			2.736***	0.176***	2.850***	0.184***			3.075***	0.188***	3.179***	0.195***			0.922	-0.00210	0.944	-0.00149
			(0.576)	(0.037)	(0.597)	(0.037)			(0.643)	(0.036)	(0.661)	(0.036)			(0.462)	(0.013)	(0.469)	(0.013)
Self-assessment=Best in class			7.147***	0.218***	6.726***	0.214***			6.910***	0.200***	6.560**	0.197***			0.277	-0.0197	0.236	-0.0210*
			(5.231)	(0.045)	(4.900)	(0.046)			(5.112)	(0.042)	(4.796)	(0.043)			(0.381)	(0.013)	(0.328)	(0.012)
Self-assessment=Among the best			5.565***	0.254***	5.622***	0.254***			4.155***	0.204***	4.174***	0.204***			0.792	-0.00580	0.730	-0.00782
			(2.621)	(0.057)	(2.680)	(0.057)			(1.928)	(0.058)	(1.942)	(0.057)			(0.810)	(0.025)	(0.760)	(0.025)
Self-assessment=Better than the majority			6.707***	0.255***	6.753***	0.255***			5.365***	0.217***	5.385***	0.217***			0.279	-0.0232	0.262	-0.0242
			(3.276)	(0.048)	(3.299)	(0.048)			(2.563)	(0.047)	(2.564)	(0.047)			(0.370)	(0.018)	(0.353)	(0.018)
Self-assessment=About as good as most people			2.787**	0.155***	2.871**	0.158***			2.258*	0.120**	2.309*	0.123**			0.867	-0.00364	0.830	-0.00477
			(1.286)	(0.061)	(1.330)	(0.061)			(0.991)	(0.059)	(1.011)	(0.059)			(0.919)	(0.027)	(0.900)	(0.028)
Reconstituted family			1.183	0.0274	1.226	0.0330			1.081	0.0122	1.114	0.0168			2.187	0.0255	2.283	0.0275
			(0.330)	(0.045)	(0.342)	(0.044)			(0.304)	(0.043)	(0.313)	(0.043)			(1.249)	(0.023)	(1.295)	(0.024)
Single parent family			0.910	-0.0159	0.912	-0.0155			0.948	-0.00842	0.954	-0.00741			0.417	-0.0169	0.449	-0.0158
			(0.274)	(0.051)	(0.275)	(0.051)			(0.291)	(0.049)	(0.293)	(0.049)			(0.391)	(0.014)	(0.407)	(0.014)
Household members			1.067	0.0107	1.037	0.00597			1.027	0.00416	1.004	0.000706			0.716	-0.00854	0.724	-0.00826
			(0.104)	(0.016)	(0.105)	(0.017)			(0.100)	(0.015)	(0.103)	(0.016)			(0.170)	(0.006)	(0.163)	(0.006)
European					1.737	0.0840					1.726	0.0782					1.749	0.0180
					(0.729)	(0.058)					(0.652)	(0.049)					(1.428)	(0.032)
non-European					5.172***	0.210***					4.345***	0.181***					2.382	0.0307
					(2.481)	(0.041)					(2.043)	(0.041)					(1.322)	(0.026)
Constant	2.370***		0.141***		0.156***		2.707***		0.224**		0.241**		0.0258***		0.189		0.202	
	(0.232)		(0.086)		(0.096)		(0.259)		(0.134)		(0.145)		(0.007)		(0.311)		(0.334)	
Observations	876	876	867	867	867	867	874	874	865	865	865	865	874	874	865	865	865	865
McFadden's Adj. R-squared	0.041	0.041	0.142	0.142	0.145	0.145	0.035	0.035	0.145	0.145	0.147	0.147	0.000	0.000	0.072	0.072	0.066	0.066
Log-likelihood	-454.718	-454.718	-479215.112	-479215.112	-477788.786	-477788.786	-436.480	-436.480	-457514.704	-457514.704	-456594.987	-456594.987	-106.321	-106.321	-108793.156	-108793.156	-109503.799	-109503.799

Standard errors in parentheses

$* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$

Note: Logistic regressions with robust standard errors clustered at the family level (all except Model 1, 4 and 7). In Models 1-3 the dependent variable is a dummy defined as having aspirations for a university degree (having responded "Yes, absolutely" or "Yes, probably" on the question "Would you like to continue going to school after the upper secondary level, that is, attend a university or university college?". The dependent variable in Models 4-6 is a dummy defined as having expectations for a university degree (having responded "Yes, absolutely" or "Yes, probably" on the question "Do you think you will actually continue going to school after the upper secondary level, that is, attend a university or university college?"). In Models 7-9 the dependent variable is a dummy for the aspirations-expectations gap which is defined as having aspirations for a university degree and but not expecting to get one (having responded "Yes, absolutely" or "Yes, probably" on the aspirations question and "No, probably not" or "No, absolutely not" on the expectations question. The reference category consists of children with at least one native-born parent.

Table A2: Odds ratios and average marginal effects from probit regressions of aspirations, expectations and the aspirations-expectations gap by region of origin, LNU

	Aspirations						Expectations						Aspirations-expectations gap					
	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7		Model 8		Model 9	
	OR	AME	OR	AME	OR	AME	OR	AME	OR	AME	OR	AME	OR	AME	OR	AME	OR	AME
Foreign-born parents	1.911***	0.178***	2.315***	0.168***			1.815***	0.157***	2.121***	0.148***			1.052	0.00312	1.753**	0.0287		
	(0.218)	(0.028)	(0.494)	(0.039)			(0.197)	(0.026)	(0.461)	(0.036)			(0.192)	(0.011)	(0.470)	(0.020)		
Age (demeaned)			0.977	-0.00657	0.975	-0.00713			0.951	-0.0136	0.950	-0.0139			0.978	-0.00130	0.982	-0.00107
			(0.031)	(0.009)	(0.031)	(0.009)			(0.031)	(0.009)	(0.031)	(0.009)			(0.061)	(0.004)	(0.060)	(0.004)
Female			2.009***	0.186***	1.904***	0.185***			2.084***	0.191***	2.006***	0.191***			0.699	-0.0283**	0.611**	-0.0278**
			(0.260)	(0.033)	(0.227)	(0.033)			(0.275)	(0.032)	(0.246)	(0.032)			(0.160)	(0.012)	(0.124)	(0.012)
Foreign-born parents × Female			0.658						0.711						0.520			
			(0.190)						(0.225)						(0.241)			
Parents' highest educ=University studies			1.800***	0.174***	1.852***	0.184***			1.923***	0.187***	1.976***	0.195***			0.978	-0.00130	0.995	-0.000307
			(0.219)	(0.037)	(0.227)	(0.037)			(0.232)	(0.036)	(0.239)	(0.036)			(0.200)	(0.012)	(0.202)	(0.012)
Self-assessment=Best in class			3.367***	0.229***	3.254***	0.226***			3.174***	0.208***	3.098***	0.206***			0.542	-0.0213*	0.508	-0.0225**
			(1.405)	(0.042)	(1.357)	(0.044)			(1.325)	(0.041)	(1.291)	(0.042)			(0.312)	(0.012)	(0.292)	(0.011)
Self-assessment=Among the best			2.884***	0.265***	2.887***	0.265***			2.345***	0.210***	2.346***	0.209***			0.902	-0.00582	0.862	-0.00834
			(0.804)	(0.057)	(0.811)	(0.057)			(0.658)	(0.060)	(0.661)	(0.059)			(0.405)	(0.025)	(0.392)	(0.025)
Self-assessment=Better than the majority			3.170***	0.264***	3.191***	0.265***			2.695***	0.222***	2.695***	0.222***			0.570	-0.0238	0.556	-0.0248
			(0.908)	(0.048)	(0.909)	(0.047)			(0.768)	(0.049)	(0.766)	(0.049)			(0.307)	(0.017)	(0.299)	(0.017)
Self-assessment=About as good as most people			1.877**	0.161***	1.911**	0.165***			1.620*	0.122**	1.644*	0.125**			0.988	-0.000693	0.958	-0.00248
			(0.512)	(0.061)	(0.521)	(0.060)			(0.433)	(0.062)	(0.439)	(0.061)			(0.444)	(0.026)	(0.444)	(0.027)
Reconstituted family			1.108	0.0286	1.139	0.0359			1.054	0.0141	1.076	0.0195			1.450	0.0268	1.484*	0.0292
			(0.180)	(0.044)	(0.185)	(0.044)			(0.171)	(0.043)	(0.175)	(0.043)			(0.348)	(0.021)	(0.353)	(0.022)
Single parent family			0.940	-0.0176	0.951	-0.0143			0.975	-0.00676	0.983	-0.00472			0.704	-0.0162	0.713	-0.0158
			(0.165)	(0.051)	(0.167)	(0.051)			(0.173)	(0.048)	(0.174)	(0.048)			(0.255)	(0.014)	(0.255)	(0.014)
Household members			1.043	0.0119	1.025	0.00690			1.021	0.00568	1.009	0.00235			0.856	-0.00899	0.862	-0.00863
			(0.057)	(0.016)	(0.058)	(0.016)			(0.056)	(0.015)	(0.058)	(0.016)			(0.082)	(0.006)	(0.079)	(0.006)
European					1.376	0.0837					1.357	0.0762					1.291	0.0181
					(0.318)	(0.055)					(0.276)	(0.046)					(0.453)	(0.030)
non-European					2.403***	0.201***					2.233***	0.176***					1.549*	0.0346
					(0.607)	(0.042)					(0.547)	(0.040)					(0.360)	(0.025)
Constant	1.705***		0.304***		0.324***		1.847***		0.407**		0.424**		0.141***		0.338		0.349	
	(0.101)		(0.108)		(0.116)		(0.105)		(0.145)		(0.151)		(0.016)		(0.232)		(0.239)	
Observations	876	876	867	867	867	867	874	874	865	865	865	865	874	874	865	865	865	865
McFadden's Adj. R-squared	0.041	0.041	0.143	0.143	0.145	0.145	0.035	0.035	0.146	0.146	0.147	0.147	0.000	0.000	0.077	0.077	0.070	0.070
Log-likelihood	-454.718	-454.718	-478698.130	-478698.130	-477793.565	-477793.565	-436.480	-436.480	-457103.955	-457103.955	-456281.039	-456281.039	-106.321	-106.321	-108173.173	-108173.173	-108968.888	-108968.888

Standard errors in parentheses

$* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$

Note: Probit regressions with robust standard errors clustered at the family level (all except Model 1, 4 and 7). In Models 1-3 the dependent variable is a dummy defined as having aspirations for a university degree (having responded "Yes, absolutely" or "Yes, probably" on the question "Would you like to continue going to school after the upper secondary level, that is, attend a university or university college?". The dependent variable in Models 4-6 is a dummy defined as having expectations for a university degree (having responded "Yes, absolutely" or "Yes, probably" on the question "Do you think you will actually continue going to school after the upper secondary level, that is, attend a university or university college?"). In Models 7-9 the dependent variable is a dummy for the aspirations-expectations gap which is defined as having aspirations for a university degree and but not expecting to get one (having responded "Yes, absolutely" or "Yes, probably" on the aspirations question and "No, probably not" or "No, absolutely not" on the expectations question. The reference category consists is children with at least one native-born parent.

Table A3: Odds ratios from random-effects probit models of aspirations, expectations and the gap in grade 8 and 9, CILS4EU

	Aspirations		Expectations		Aspirations-expectations gap	
	Grade 8	Grade 9	Grade 8	Grade 9	Grade 8	Grade 9
Foreign-born parents	2.221***	2.283***	1.956***	2.213***	1.080	0.961
	(0.173)	(0.214)	(0.145)	(0.199)	(0.087)	(0.097)
Female	1.348***	1.553***	1.206***	1.519***	1.116*	1.037
	(0.072)	(0.100)	(0.063)	(0.095)	(0.065)	(0.075)
Foreign-born parents × Female	0.883	0.892	0.857*	0.843	0.951	0.875
	(0.088)	(0.113)	(0.079)	(0.098)	(0.097)	(0.115)
Higher educated parents	1.433***	1.347***	1.391***	1.515***	0.939	0.790***
	(0.066)	(0.076)	(0.061)	(0.082)	(0.046)	(0.050)
Language test score (std)	1.415***	1.144***	1.397***	1.127***	0.974	0.992
	(0.042)	(0.043)	(0.040)	(0.041)	(0.030)	(0.041)
Cognitive test score (std)	1.126***	1.052	1.193***	1.041	0.931**	0.996
	(0.030)	(0.035)	(0.031)	(0.034)	(0.026)	(0.037)
All friends: female	0.920	1.036	1.046	1.032	0.866**	0.997
	(0.056)	(0.079)	(0.062)	(0.075)	(0.057)	(0.080)
All friends: foreign-born parents	1.225***	1.046	1.148*	0.984	1.033	1.075
	(0.096)	(0.099)	(0.084)	(0.089)	(0.079)	(0.103)
All friends: higher educ parents	1.012	1.018	0.996	0.884	0.938	1.155
	(0.072)	(0.089)	(0.068)	(0.074)	(0.071)	(0.104)
All friends: university aspirations	1.254***	1.132*	1.196***	1.148**	1.013	0.931
	(0.072)	(0.074)	(0.063)	(0.072)	(0.055)	(0.064)
Grade sum		1.332***		1.379***		0.892***
		(0.024)		(0.024)		(0.017)
Observations	4075	3381	4075	3377	4075	3377
Log-likelihood	−2285.864	−1473.354	−2539.672	−1637.783	−1888.582	−1096.194

Exponentiated coefficients; Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Random-effects probit regressions estimated using the xtprobit command in Stata.

Table A4: Odds ratios from random-effects logistic models of aspirations in grade 8 and 9, CILS4EU

	Aspirations grade 8			Aspirations grade 9			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Foreign-born parents	1.907***	2.001***	3.858***	1.449***	2.178***	4.428***	4.188***
	(0.160)	(0.177)	(0.514)	(0.114)	(0.239)	(0.707)	(0.694)
Female			1.638***			2.203***	2.124***
			(0.145)			(0.238)	(0.237)
Foreign-born parents × Female			0.820			0.833	0.841
			(0.141)			(0.190)	(0.192)
Higher educated parents			1.837***			1.674***	1.663***
			(0.143)			(0.167)	(0.166)
Language test score (std)			1.785***			1.222***	1.228***
			(0.090)			(0.080)	(0.080)
Cognitive test score (std)			1.218***			1.075	1.079
			(0.054)			(0.062)	(0.062)
All friends: female			0.875				1.094
			(0.091)				(0.148)
All friends: foreign-born parents			1.393**				1.095
			(0.187)				(0.186)
All friends: higher educ parents			1.024				1.056
			(0.123)				(0.164)
All friends: university aspirations			1.469***				1.258**
			(0.143)				(0.145)
Grade sum						1.707***	1.698***
						(0.057)	(0.057)
Observations	4364	4075	4075	4354	3381	3381	3381
Log-likelihood	−2704.027	−2500.710	−2285.852	−2833.177	−1816.393	−1470.843	−1467.634

Exponentiated coefficients; Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Random-effects logit regressions estimated using the xtlogit command in Stata.

Table A5: Odds ratios from random-effects logistic models of expectations in grade 8 and 9, CILS4EU

	Expectations grade 8			Expectations grade 9			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Foreign-born parents	1.451***	1.489***	3.039***	1.414***	1.839***	4.030***	3.933***
	(0.111)	(0.118)	(0.375)	(0.108)	(0.180)	(0.609)	(0.614)
Female			1.363***			2.088***	2.040***
			(0.116)			(0.217)	(0.220)
Foreign-born parents × Female			0.778			0.776	0.778
			(0.119)			(0.158)	(0.159)
Higher educated parents			1.715***			2.042***	2.046***
			(0.123)			(0.191)	(0.192)
Language test score (std)			1.738***			1.203***	1.205***
			(0.084)			(0.075)	(0.076)
Cognitive test score (std)			1.340***			1.062	1.068
			(0.057)			(0.059)	(0.059)
All friends: female			1.078				1.063
			(0.105)				(0.135)
All friends: foreign-born parents			1.250*				0.990
			(0.151)				(0.156)
All friends: higher educ parents			0.994				0.832
			(0.112)				(0.122)
All friends: university aspirations			1.347***				1.301**
			(0.116)				(0.142)
Grade sum						1.781***	1.777***
						(0.057)	(0.057)
Observations	4364	4075	4075	4353	3377	3377	3377
Log-likelihood	−2962.101	−2764.690	−2539.116	−2945.748	−2068.819	−1633.198	−1629.471

Exponentiated coefficients; Standard errors in parentheses

$* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$

Note: Random-effects logit regressions estimated using the xtlogit command in Stata.

Table A6: Odds ratios from random-effects logistic models of the aspirations-expectations gap in grade 8 and 9, CILS4EU

	Aspirations-expectations gap grade 8			Aspirations-expectations gap grade 9			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Foreign-born parents	1.246**	1.238*	1.149	0.970	0.998	0.966	0.941
	(0.109)	(0.141)	(0.164)	(0.114)	(0.124)	(0.176)	(0.181)
Female			1.215*			1.082	1.077
			(0.127)			(0.146)	(0.150)
Foreign-born parents × Female			0.918			0.761	0.762
			(0.164)			(0.191)	(0.192)
Higher educated parents			0.898			0.637***	0.628***
			(0.078)			(0.079)	(0.078)
Language test score (std)			0.954			0.985	0.990
			(0.052)			(0.078)	(0.078)
Cognitive test score (std)			0.883**			0.987	0.983
			(0.044)			(0.069)	(0.069)
All friends: female			0.776**				1.003
			(0.091)				(0.154)
All friends: foreign-born parents			1.056				1.153
			(0.143)				(0.211)
All friends: higher educ parents			0.894				1.305
			(0.123)				(0.224)
All friends: university aspirations			1.024				0.868
			(0.100)				(0.116)
Grade sum						0.807***	0.808***
						(0.030)	(0.030)
Observations	4364	3796	4075	4349	3377	3377	3377
Log-likelihood	−2033.951	−1422.293	−1888.730	−1316.800	−1139.609	−1098.698	−1097.010

Exponentiated coefficients; Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Random-effects logit regressions estimated using the xtlogit command in Stata.

Table A7: Odds ratios from fixed-effects logistic models of aspirations in grade 8 and 9, CILS4EU

	Aspirations grade 8			Aspirations grade 9			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Foreign-born parents	1.695***	1.804***	3.755***	1.551***	2.348***	4.476***	4.459***
	(0.162)	(0.182)	(0.533)	(0.143)	(0.308)	(0.816)	(0.815)
Female			1.716***			2.103***	2.049***
			(0.157)			(0.241)	(0.243)
Foreign-born parents × Female			0.746*			0.774	0.753
			(0.133)			(0.185)	(0.180)
Higher educated parents			1.739***			1.537***	1.531***
			(0.140)			(0.161)	(0.161)
Language test score (std)			1.838***			1.252***	1.258***
			(0.096)			(0.088)	(0.088)
Cognitive test score (std)			1.211***			1.061	1.062
			(0.056)			(0.067)	(0.067)
All friends: female			0.940				1.178
			(0.112)				(0.191)
All friends: foreign-born parents			1.307				1.511
			(0.248)				(0.398)
All friends: higher educ parents			1.092				0.959
			(0.155)				(0.186)
All friends: university aspirations			0.814*				0.870
			(0.089)				(0.129)
Grade sum						1.683***	1.686***
						(0.060)	(0.060)
Observations	4308	4003	4003	4312	3177	3177	3177
Log-likelihood	−2121.483	−1932.434	−1746.496	−2241.338	−1324.658	−1035.668	−1033.601

Exponentiated coefficients; Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Fixed-effects logistic regressions estimated using xtlogit in Stata.

Table A8: Odds ratios from fixed-effects logistic models of expectations in grade 8 and 9, CILS4EU

	Expectations grade 8			Expectations grade 9			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Foreign-born parents	1.354***	1.405***	2.881***	1.528***	1.998***	4.266***	4.253***
	(0.118)	(0.128)	(0.375)	(0.136)	(0.230)	(0.726)	(0.725)
Female			1.407***			1.993***	1.947***
			(0.124)			(0.218)	(0.220)
Foreign-born parents × Female			0.736*			0.710	0.703*
			(0.116)			(0.151)	(0.150)
Higher educated parents			1.599***			1.862***	1.852***
			(0.118)			(0.181)	(0.180)
Language test score (std)			1.784***			1.222***	1.224***
			(0.089)			(0.082)	(0.082)
Cognitive test score (std)			1.316***			1.051	1.053
			(0.059)			(0.063)	(0.063)
All friends: female			1.135				1.121
			(0.126)				(0.166)
All friends: foreign-born parents			1.196				1.226
			(0.201)				(0.283)
All friends: higher educ parents			0.893				0.689**
			(0.117)				(0.124)
All friends: university aspirations			0.959				1.170
			(0.097)				(0.162)
Grade sum						1.769***	1.776***
						(0.060)	(0.060)
Observations	4344	4055	4055	4311	3298	3298	3298
Log-likelihood	−2349.495	−2168.771	−1975.535	−2342.992	−1536.249	−1164.510	−1161.053

Exponentiated coefficients; Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Fixed-effects logistic regressions estimated using xtlogit in Stata.

Table A9: Odds ratios from fixed-effects logistic models of the aspirations-expectations gap in grade 8 and 9, CILS4EU

	Aspirations-expectations gap grade 8			Aspirations-expectations gap grade 9			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Foreign-born parents	1.235*	1.238*	1.193	0.940	0.936	0.901	0.896
	(0.135)	(0.141)	(0.184)	(0.140)	(0.150)	(0.190)	(0.189)
Female			1.218*			1.103	1.100
			(0.132)			(0.157)	(0.160)
Foreign-born parents × Female			0.918			0.785	0.778
			(0.172)			(0.209)	(0.207)
Higher educated parents			0.945			0.640***	0.643***
			(0.085)			(0.084)	(0.085)
Language test score (std)			0.939			0.984	0.988
			(0.053)			(0.082)	(0.082)
Cognitive test score (std)			0.906*			1.012	1.012
			(0.047)			(0.076)	(0.076)
All friends: female			0.800				1.044
			(0.110)				(0.195)
All friends: foreign-born parents			1.113				1.188
			(0.228)				(0.328)
All friends: higher educ parents			1.105				1.448*
			(0.179)				(0.317)
All friends: university aspirations			0.869				0.743*
			(0.106)				(0.132)
Grade sum						0.802***	0.801***
						(0.033)	(0.033)
Observations	4106	3796	3796	3373	2630	2630	2630
Log-likelihood	−1551.434	−1422.293	−1414.652	−930.517	−783.482	−750.332	−747.804

Exponentiated coefficients; Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Fixed-effects logistic regressions estimated using xtlogit in Stata.

Table A10: Odds ratios and average marginal effects from logistic regressions of aspirations, expectations and the aspirations-expectations gap by region of origin, LNU

	Aspirations						Expectations						Aspirations-expectations gap					
	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7		Model 8		Model 9	
	OR	AME	OR	AME	OR	AME	OR	AME	OR	AME	OR	AME	OR	AME	OR	AME	OR	AME
Foreign-born parents	3.132***	0.178***	4.361***	0.173***			2.894***	0.157***	3.821***	0.152***			1.128	0.00312	3.009*	0.0252		
	(0.657)	(0.028)	(1.754)	(0.040)			(0.582)	(0.026)	(1.553)	(0.038)			(0.488)	(0.011)	(1.803)	(0.021)		
Age (demeaned)			0.962	-0.00650	0.961	-0.00669			0.922	-0.0128	0.921	-0.0129			0.952	-0.00125	0.954	-0.00120
			(0.054)	(0.009)	(0.053)	(0.009)			(0.054)	(0.009)	(0.053)	(0.009)			(0.149)	(0.004)	(0.148)	(0.004)
Female			3.192***	0.184***	3.022***	0.186***			3.459***	0.188***	3.301***	0.190***			0.458	-0.0266**	0.339**	-0.0265**
			(0.706)	(0.033)	(0.635)	(0.034)			(0.798)	(0.032)	(0.719)	(0.033)			(0.256)	(0.013)	(0.172)	(0.013)
Foreign-born parents × Female			0.515						0.561						0.238			
			(0.282)						(0.339)						(0.295)			
Parents' highest educ=University studies			2.736***	0.176***	2.850***	0.184***			3.075***	0.188***	3.179***	0.195***			0.922	-0.00210	0.944	-0.00149
			(0.576)	(0.037)	(0.597)	(0.037)			(0.643)	(0.036)	(0.661)	(0.036)			(0.462)	(0.013)	(0.469)	(0.013)
Self-assessment=Best in class			7.147***	0.218***	6.726***	0.214***			6.910***	0.200***	6.560**	0.197***			0.277	-0.0197	0.236	-0.0210*
			(5.231)	(0.045)	(4.900)	(0.046)			(5.112)	(0.042)	(4.796)	(0.043)			(0.381)	(0.013)	(0.328)	(0.012)
Self-assessment=Among the best			5.565***	0.254***	5.622***	0.254***			4.155***	0.204***	4.174***	0.204***			0.792	-0.00580	0.730	-0.00782
			(2.621)	(0.057)	(2.680)	(0.057)			(1.928)	(0.058)	(1.942)	(0.057)			(0.810)	(0.025)	(0.760)	(0.025)
Self-assessment=Better than the majority			6.707***	0.255***	6.753***	0.255***			5.365***	0.217***	5.385***	0.217***			0.279	-0.0232	0.262	-0.0242
			(3.276)	(0.048)	(3.299)	(0.048)			(2.563)	(0.047)	(2.564)	(0.047)			(0.370)	(0.018)	(0.353)	(0.018)
Self-assessment=About as good as most people			2.787**	0.155***	2.871**	0.158***			2.258*	0.120**	2.309*	0.123**			0.867	-0.00364	0.830	-0.00477
			(1.286)	(0.061)	(1.330)	(0.061)			(0.991)	(0.059)	(1.011)	(0.059)			(0.919)	(0.027)	(0.900)	(0.028)
Reconstituted family			1.183	0.0274	1.226	0.0330			1.081	0.0122	1.114	0.0168			2.187	0.0255	2.283	0.0275
			(0.330)	(0.045)	(0.342)	(0.044)			(0.304)	(0.043)	(0.313)	(0.043)			(1.249)	(0.023)	(1.295)	(0.024)
Single parent family			0.910	-0.0159	0.912	-0.0155			0.948	-0.00842	0.954	-0.00741			0.417	-0.0169	0.449	-0.0158
			(0.274)	(0.051)	(0.275)	(0.051)			(0.291)	(0.049)	(0.293)	(0.049)			(0.391)	(0.014)	(0.407)	(0.014)
Household members			1.067	0.0107	1.037	0.00597			1.027	0.00416	1.004	0.000706			0.716	-0.00854	0.724	-0.00826
			(0.104)	(0.016)	(0.105)	(0.017)			(0.100)	(0.015)	(0.103)	(0.016)			(0.170)	(0.006)	(0.163)	(0.006)
European					1.737	0.0840					1.726	0.0782					1.749	0.0180
					(0.729)	(0.058)					(0.652)	(0.049)					(1.428)	(0.032)
non-European					5.172***	0.210***					4.345***	0.181***					2.382	0.0307
					(2.481)	(0.041)					(2.043)	(0.041)					(1.322)	(0.026)
Constant	2.370***		0.141***		0.156***		2.707***		0.224**		0.241**		0.0258***		0.189		0.202	
	(0.232)		(0.086)		(0.096)		(0.259)		(0.134)		(0.145)		(0.007)		(0.311)		(0.334)	
Observations	876	876	867	867	867	867	874	874	865	865	865	865	874	874	865	865	865	865
McFadden's Adj. R-squared	0.041	0.041	0.142	0.142	0.145	0.145	0.035	0.035	0.145	0.145	0.147	0.147	0.000	0.000	0.072	0.072	0.066	0.066
Log-likelihood	-454.718	-454.718	-479215.112	-479215.112	-477788.786	-477788.786	-436.480	-436.480	-457514.704	-457514.704	-456594.987	-456594.987	-106.321	-106.321	-108793.156	-108793.156	-109503.799	-109503.799

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Logistic regressions with robust standard errors clustered at the family level (all except Model 1, 4 and 7). In Models 1-3 the dependent variable is a dummy defined as having aspirations for a university degree (having responded "Yes, absolutely" or "Yes, probably" on the question "Would you like to continue going to school after the upper secondary level, that is, attend a university or university college?". The dependent variable in Models 4-6 is a dummy defined as having expectations for a university degree (having responded "Yes, absolutely" or "Yes, probably" on the question "Do you think you will actually continue going to school after the upper secondary level, that is, attend a university or university college?"). In Models 7-9 the dependent variable is a dummy for the aspirations-expectations gap which is defined as having aspirations for a university degree and but not expecting to get one (having responded "Yes, absolutely" or "Yes, probably" on the aspirations question and "No, probably not" or "No, absolutely not" on the expectations question. The reference category consists of children with at least one native-born parent.

Table A11: Log odds and odds ratios from multileveled random-effects logistic regressions of aspirations in grade 8 and 9 respectively, CILS4EU

	Aspirations grade 8						Aspirations grade 9							
	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
	logodds	OR	logodds	OR	logodds	OR	logodds	OR	logodds	OR	logodds	OR	logodds	OR
Foreign-born parents	0.645***	1.907***	0.693***	2.001***	1.350***	3.858***	0.371***	1.449***	0.778***	2.178***	1.488***	4.428***	1.432***	4.188***
	(0.090)	(0.171)	(0.099)	(0.198)	(0.143)	(0.553)	(0.078)	(0.113)	(0.111)	(0.243)	(0.172)	(0.761)	(0.173)	(0.727)
Female					0.493***	1.638***					0.790***	2.203***	0.753***	2.124***
					(0.102)	(0.166)					(0.122)	(0.268)	(0.123)	(0.262)
Foreign-born parents × Female					-0.198	0.820					-0.183	0.833	-0.173	0.841
					(0.187)	(0.153)					(0.240)	(0.200)	(0.240)	(0.202)
Higher educated parents					0.608***	1.837***					0.515***	1.674***	0.509***	1.663***
					(0.080)	(0.147)					(0.097)	(0.163)	(0.097)	(0.161)
Language test score (std)					0.580***	1.785***					0.201***	1.222***	0.206***	1.228***
					(0.054)	(0.097)					(0.066)	(0.081)	(0.066)	(0.081)
Cognitive test score (std)					0.197***	1.218***					0.0720	1.075	0.0756	1.079
					(0.048)	(0.059)					(0.060)	(0.065)	(0.060)	(0.065)
All friends: female					-0.133	0.875							0.0896	1.094
					(0.111)	(0.097)							(0.126)	(0.138)
All friends: foreign-born parents					0.332**	1.393**							0.0908	1.095
					(0.147)	(0.204)							(0.179)	(0.196)
All friends: higher educ parents					0.0238	1.024							0.0540	1.056
					(0.121)	(0.124)							(0.152)	(0.161)
All friends: university aspirations					0.385***	1.469***							0.230**	1.258**
					(0.114)	(0.168)							(0.111)	(0.139)
Grade sum											0.534***	1.707***	0.529***	1.698***
											(0.041)	(0.070)	(0.041)	(0.069)
Constant	0.551***	1.735***	0.583***	1.791***	-0.258***	0.773***	0.432***	1.540***	1.016***	2.761***	-2.125***	0.119***	-2.174***	0.114***
	(0.055)	(0.096)	(0.057)	(0.101)	(0.086)	(0.066)	(0.050)	(0.077)	(0.062)	(0.170)	(0.176)	(0.021)	(0.179)	(0.020)
var(_cons[classid])														
Constant	0.239***	1.271***	0.256***	1.292***	0.143***	1.154***	0.204***	1.227***	0.239***	1.270***	0.114*	1.121*	0.0901	1.094
	(0.052)	(0.067)	(0.056)	(0.072)	(0.052)	(0.060)	(0.065)	(0.080)	(0.060)	(0.076)	(0.059)	(0.067)	(0.056)	(0.062)
Observations	4364	4364	4075	4075	4075	4075	4354	4354	3381	3381	3381	3381	3381	3381
Log-likelihood	-2704.027	-2704.027	-2500.710	-2500.710	-2285.852	-2285.852	-2833.177	-2833.177	-1816.393	-1816.393	-1470.843	-1470.843	-1467.634	-1467.634

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Multilevel logistic regressions. Standard errors are clustered at the classroom level. All regressions include a random effect estimate. Variable definitions are found in section 3.7. The dependent variable is a dummy defined as having aspirations for a university degree ("What is the highest level of education you wish to get? (Don't know, No degree, Compulsory school, Upper secondary school, College/university)").

Table A12: Log odds and odds ratios from multileveled random-effects logistic regressions of expectations in grade 8 and 9 respectively, CILS4EU

| | Expectations grade 8 | | | | | | Expectations grade 9 | | | | | | | |
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | | Model 6 | | Model 7 | |
	logodds	OR	logodds	OR	logodds	OR	logodds	OR	logodds	OR	logodds	OR	logodds	OR
Foreign-born parents	0.372***	1.451***	0.398***	1.489***	1.112***	3.039***	0.347***	1.414***	0.609***	1.839***	1.394***	4.030***	1.369***	3.933***
	(0.075)	(0.108)	(0.081)	(0.121)	(0.130)	(0.394)	(0.077)	(0.109)	(0.100)	(0.183)	(0.159)	(0.639)	(0.159)	(0.625)
Female					0.310***	1.363***					0.736***	2.088***	0.713***	2.040***
					(0.096)	(0.131)					(0.109)	(0.229)	(0.112)	(0.228)
Foreign-born parents × Female					-0.251	0.778					-0.254	0.776	-0.250	0.778
					(0.158)	(0.123)					(0.213)	(0.165)	(0.212)	(0.165)
Higher educated parents					0.539***	1.715***					0.714***	2.042***	0.716***	2.046***
					(0.077)	(0.132)					(0.093)	(0.191)	(0.093)	(0.190)
Language test score (std)					0.553***	1.738***					0.185***	1.203***	0.186***	1.205***
					(0.050)	(0.087)					(0.062)	(0.075)	(0.062)	(0.075)
Cognitive test score (std)					0.292***	1.340***					0.0604	1.062	0.0658	1.068
					(0.043)	(0.058)					(0.052)	(0.055)	(0.052)	(0.056)
All friends: female					0.0750	1.078							0.0607	1.063
					(0.115)	(0.123)							(0.116)	(0.123)
All friends: foreign-born parents					0.223*	1.250*							-0.0104	0.990
					(0.129)	(0.162)							(0.172)	(0.170)
All friends: higher educ parents					-0.00636	0.994							-0.183	0.832
					(0.121)	(0.121)							(0.152)	(0.127)
All friends: university aspirations					0.298***	1.347***							0.263**	1.301**
					(0.098)	(0.132)							(0.109)	(0.141)
Grade sum											0.577***	1.781***	0.575***	1.777***
											(0.039)	(0.070)	(0.039)	(0.070)
Constant	-0.0270	0.973	-0.00201	0.998	-0.791***	0.453***	0.137***	1.147***	0.634***	1.886***	-2.861***	0.0572***	-2.899***	0.0551***
	(0.055)	(0.053)	(0.055)	(0.055)	(0.086)	(0.039)	(0.052)	(0.060)	(0.061)	(0.114)	(0.192)	(0.011)	(0.195)	(0.011)
var(_cons[classid])														
Constant	0.263***	1.301***	0.256***	1.292***	0.173***	1.189***	0.211***	1.236***	0.283***	1.327***	0.179***	1.197***	0.172***	1.187***
	(0.048)	(0.062)	(0.047)	(0.061)	(0.044)	(0.052)	(0.056)	(0.070)	(0.063)	(0.084)	(0.065)	(0.077)	(0.064)	(0.076)
Observations	4364	4364	4075	4075	4075	4075	4353	4353	3377	3377	3377	3377	3377	3377
Log-likelihood	-2962.101	-2962.101	-2764.690	-2764.690	-2539.116	-2539.116	-2945.748	-2945.748	-2068.819	-2068.819	-1633.198	-1633.198	-1629.471	-1629.471

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Multilevel logistic regression. Standard errors are clustered at the classroom level. All regressions include a random effect estimate. Variable definitions are found in section 3.7. The dependent variable is a dummy defined as having expectations for a university degree ("What is the highest level of education think you will actually get? (Don't know, No degree, Compulsory school, Upper secondary school, College/university)").

Table A13: Log odds coefficients and odds ratios from regressions predicting the aspirations-expectations gap in grade 8 and 9 respectively, CILS4EU

| | Aspirations-expectations gap grade 8 | | | | | | Aspirations-expectations gap grade 9 | | | | | | | |
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | | Model 6 | | Model 7 | |
	logodds	OR	logodds	OR	logodds	OR	logodds	OR	logodds	OR	logodds	OR	logodds	OR
Foreign-born parents	0.220***	1.246***	0.223**	1.250**	0.0940	1.099	-0.0307	0.970	-0.00185	0.998	-0.169	0.845	-0.192	0.825
	(0.084)	(0.105)	(0.091)	(0.114)	(0.104)	(0.114)	(0.116)	(0.112)	(0.124)	(0.124)	(0.136)	(0.115)	(0.151)	(0.124)
Female					0.167*	1.182*					0.000870	1.001	-0.00424	0.996
					(0.092)	(0.108)					(0.114)	(0.114)	(0.119)	(0.119)
Higher educated parents					-0.108	0.897					-0.453***	0.636***	-0.467***	0.627***
					(0.086)	(0.077)					(0.124)	(0.079)	(0.125)	(0.078)
Language test score (std)					-0.0464	0.955					-0.0135	0.987	-0.00897	0.991
					(0.054)	(0.052)					(0.079)	(0.078)	(0.079)	(0.079)
Cognitive test score (std)					-0.125***	0.882***					-0.0116	0.988	-0.0159	0.984
					(0.047)	(0.042)					(0.070)	(0.069)	(0.070)	(0.069)
All friends: female					-0.253*	0.777*							0.00541	1.005
					(0.131)	(0.102)							(0.154)	(0.155)
All friends: foreign-born parents					0.0517	1.053							0.135	1.145
					(0.154)	(0.163)							(0.183)	(0.210)
All friends: higher educ parents					-0.111	0.895							0.271	1.312
					(0.131)	(0.117)							(0.171)	(0.225)
All friends: university aspirations					0.0241	1.024							-0.141	0.869
					(0.099)	(0.101)							(0.133)	(0.116)
Grade sum											-0.214***	0.807***	-0.213***	0.808***
											(0.037)	(0.030)	(0.037)	(0.030)
Constant	-1.630***	0.196***	-1.635***	0.195***	-1.580***	0.206***	-2.322***	0.0981***	-2.135***	0.118***	-0.930***	0.395***	-0.933***	0.393***
	(0.055)	(0.011)	(0.058)	(0.011)	(0.087)	(0.018)	(0.074)	(0.007)	(0.066)	(0.008)	(0.173)	(0.068)	(0.174)	(0.068)
var(_cons[classid])														
Constant	0.0823*	1.086*	0.116**	1.123**	0.0993*	1.104*	0.0478	1.049						
	(0.046)	(0.050)	(0.053)	(0.059)	(0.053)	(0.058)	(0.068)	(0.072)						
Observations	4364	4364	4075	4075	4075	4075	4349	4349	3377	3377	3377	3377	3377	3377
Log-likelihood	-2033.951	-2033.951	-1899.488	-1899.488	-1888.844	-1888.844	-1316.800	-1316.800	-1139.609	-1139.609	-1099.288	-1099.288	-1097.592	-1097.592

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Multilevel logistic regression. Standard errors are clustered at the classroom level for all except models (4)-(6). All except models (4)-(6) include a random effect estimate. Variable definitions are found in section 3.7. The dependent variable is a dummy defined as having aspirations for a university degree ("What is the highest level of education you wish to get? (Don't know, No degree, Compulsory school, Upper secondary school, College/university)") and not expecting to get one ("What is the highest level of education you think you will actually get? (Don't know, No degree, Compulsory school, Upper secondary school, College/university)").

Table B1: Ordered logit and probit regressions, LNU 2010

	Aspirations		Expectations	
	Ordered logit	Ordered probit	Ordered logit	Ordered probit
	(1)	(2)	(3)	(4)
Foreign-born parents	1.127***	0.638***	0.938***	0.546***
	(0.252)	(0.141)	(0.205)	(0.118)
Age (demeaned)	0.026	0.013	0.066	0.027
	(0.044)	(0.025)	(0.044)	(0.025)
Female	0.867***	0.538***	1.046***	0.634***
	(0.160)	(0.093)	(0.164)	(0.094)
Parents' highest educ=University studies	0.854***	0.489***	0.918***	0.515***
	(0.177)	(0.101)	(0.179)	(0.101)
Self-assessment=Best in class	2.995***	1.769***	3.035***	1.819***
	(0.693)	(0.374)	(0.621)	(0.339)
Self-assessment=Among the best	1.929***	1.142***	1.911***	1.151***
	(0.389)	(0.224)	(0.335)	(0.196)
Self-assessment=Better than the majority	1.915***	1.145***	1.819***	1.133***
	(0.399)	(0.231)	(0.328)	(0.194)
Self-assessment=About as good as most people	1.140***	0.683***	1.167***	0.740***
	(0.382)	(0.220)	(0.322)	(0.189)
Reconstituted family	0.093	0.035	0.128	0.072
	(0.217)	(0.126)	(0.217)	(0.126)
Single parent family	-0.071	-0.068	-0.057	-0.028
	(0.252)	(0.151)	(0.275)	(0.157)
Household members	0.008	0.011	-0.050	-0.032
	(0.077)	(0.046)	(0.074)	(0.044)
Observations	867	867	865	865
McFadden's Adj. R-squared	0.0871	0.0881	0.0898	0.0907

Robust standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Note: Coeffients from ordered logit and probit regressions. Robust standard errors clustered at the family level.

Table B2: Bivariate ordered probit regression, LNU 2010

	Aspirations
Aspirations	
Foreign-born parents	0.621***
	(5.83)
Expectations	
Foreign-born parents	0.520***
	(5.20)
athrho	
Constant	1.499***
	(17.14)
cut11	
Constant	-1.270***
	(-17.35)
cut12	
Constant	-0.490***
	(-8.68)
cut13	
Constant	0.333***
	(6.11)
cut21	
Constant	-1.390***
	(-17.52)
cut22	
Constant	-0.586***
	(-10.08)
cut23	
Constant	0.423***
	(7.75)
Observations	874

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B3: Predicted joint probabilities from bivariate ordered regressions (SUR) of aspirations and expectations with immigrant dummy, children with native-born parent(s)

	Aspirations			
Expectations	No, absolutely not	No, probably not	Yes, probably	Yes, absolutely
No, absolutely not	0.0626	0.0189	0.0008	0.0000
No, probably not	0.0366	0.1155	0.0433	0.0011
Yes, probably	0.0028	0.0744	0.2258	0.0820
Yes, absolutely	0.0000	0.0012	0.0485	0.2864

Table B4: Predicted joint probabilities from bivariate ordered regressions (SUR) of aspirations and expectations with immigrant dummy, children with foreign-born parents

	Aspirations			
Expectations	No, absolutely not	No, probably not	Yes, probably	Yes, absolutely
No, absolutely not	0.0175	0.0099	0.0007	0.0000
No, probably not	0.0110	0.0580	0.0357	0.0016
Yes, probably	0.0008	0.0356	0.1800	0.1108
Yes, absolutely	0.0000	0.0005	0.0370	0.5009

Table C1: OLS coefficients from regressions of aspirations in grade 8 and 9 respectively, CILS4EU

	Aspirations grade 8			Aspirations grade 9			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Born abroad	−0.000829	0.0120	0.122***	−0.00392	0.0736***	0.174***	0.174***
	(0.025)	(0.026)	(0.034)	(0.026)	(0.026)	(0.041)	(0.040)
Female			0.0952***			0.107***	0.101***
			(0.019)			(0.017)	(0.018)
Born abroad × Female			−0.0807*			−0.0963**	−0.0985**
			(0.045)			(0.047)	(0.047)
Higher educated parents			0.0979***			0.0588***	0.0584***
			(0.016)			(0.014)	(0.014)
Language test score (std)			0.0954***			0.0128	0.0140
			(0.010)			(0.010)	(0.010)
Cognitive test score (std)			0.0391***			0.0116	0.0114
			(0.010)			(0.010)	(0.010)
All friends: female			−0.0109				0.0261
			(0.025)				(0.020)
All friends: foreign-born parents			0.0585*				0.0585*
			(0.035)				(0.032)
All friends: higher educ parents			0.00976				−0.0117
			(0.027)				(0.026)
All friends: university aspirations			−0.0368*				−0.0228
			(0.021)				(0.019)
Grade sum						0.0754***	0.0756***
						(0.005)	(0.005)
Constant	0.667***	0.673***	0.564***	0.630***	0.754***	0.293***	0.290***
	(0.003)	(0.003)	(0.014)	(0.003)	(0.002)	(0.025)	(0.025)
Observations	4364	4075	4075	4354	3381	3381	3381
Adjusted R^2	−0.000	−0.000	0.075	−0.000	0.002	0.164	0.164

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: OLS regressions. Standard errors are clustered at the classroom level. All regressions include classroom fixed-effects. Variable definitions are found in section 3.7. The dependent variable is a dummy defined as having aspirations for a university degree ("What is the highest level of education you wish to get? (Don't know, No degree, Compulsory school, Upper secondary school, College/university)").

Table C2: OLS coefficients from regressions of expectations in grade 8 and 9 respectively, CILS4EU

	Expectations grade 8			Expectations grade 9			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Born abroad	−0.00181	0.00208	0.0975***	−0.00946	0.0477*	0.143***	0.140***
	(0.023)	(0.025)	(0.037)	(0.026)	(0.028)	(0.036)	(0.036)
Female			0.0559***			0.104***	0.0978***
			(0.020)			(0.017)	(0.018)
Born abroad × Female			−0.0369			−0.0673	−0.0679
			(0.050)			(0.049)	(0.049)
Higher educated parents			0.0953***			0.0995***	0.0985***
			(0.017)			(0.015)	(0.015)
Language test score (std)			0.105***			0.00847	0.00846
			(0.009)			(0.011)	(0.011)
Cognitive test score (std)			0.0585***			0.0104	0.0107
			(0.009)			(0.009)	(0.009)
All friends: female			0.0279				0.0242
			(0.028)				(0.021)
All friends: foreign-born parents			0.0497				0.0372
			(0.040)				(0.041)
All friends: higher educ parents			−0.0273				−0.0578**
			(0.030)				(0.029)
All friends: university aspirations			−0.00893				0.0147
			(0.024)				(0.020)
Grade sum						0.0933***	0.0936***
						(0.005)	(0.005)
Constant	0.522***	0.528***	0.427***	0.562***	0.676***	0.110***	0.105***
	(0.002)	(0.003)	(0.014)	(0.003)	(0.003)	(0.026)	(0.027)
Observations	4364	4075	4075	4353	3377	3377	3377
Adjusted R^2	−0.000	−0.000	0.080	−0.000	0.001	0.203	0.203

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: OLS regressions. Standard errors are clustered at the classroom level. All regressions include classroom fixed-effects. Variable definitions are found in section 3.7. The dependent variable is a dummy defined as having expectations for a university degree ("What is the highest level of education think you will actually get? (Don't know, No degree, Compulsory school, Upper secondary school, College/university)").

Table C3: OLS coefficients from regressions predicting the aspirations-expectations gap in grade 8 and 9 respectively, CILS4EU

	Aspirations-expectations gap grade 8			Aspirations-expectations gap grade 9			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Born abroad	0.0190	0.0256	0.0413	−0.00115	0.0153	0.0269	0.0289
	(0.020)	(0.022)	(0.035)	(0.015)	(0.021)	(0.029)	(0.029)
Female			0.0298**			0.00845	0.00868
			(0.014)			(0.012)	(0.012)
Born abroad × Female			−0.0485			−0.0461	−0.0473
			(0.044)			(0.039)	(0.039)
Higher educated parents			−0.00817			−0.0402***	−0.0397***
			(0.013)			(0.012)	(0.012)
Language test score (std)			−0.0106			0.000640	0.00154
			(0.008)			(0.008)	(0.008)
Cognitive test score (std)			−0.0147**			−0.000246	−0.000589
			(0.007)			(0.007)	(0.007)
All friends: female			−0.0323				0.00281
			(0.021)				(0.017)
All friends: foreign-born parents			0.0193				0.0146
			(0.035)				(0.028)
All friends: higher educ parents			0.0134				0.0349*
			(0.021)				(0.021)
All friends: university aspirations			−0.0204				−0.0272**
			(0.020)				(0.014)
Grade sum						−0.0190***	−0.0190***
						(0.003)	(0.003)
Constant	0.175***	0.175***	0.175***	0.0903***	0.104***	0.213***	0.214***
	(0.002)	(0.002)	(0.011)	(0.002)	(0.002)	(0.017)	(0.018)
Observations	4364	4075	4075	4349	3377	3377	3377
Adjusted R^2	−0.000	0.000	0.003	−0.000	−0.000	0.018	0.018

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: OLS regressions. Standard errors are clustered at the classroom level. All models include classroom fixed-effects. Variable definitions are found in section 3.7. The dependent variable is a dummy defined as having aspirations for a university degree (having responded "College/university" on the question "What is the highest level of education you wish to get? (Don't know, No degree, Compulsory school, Upper secondary school, College/university)" but not expecting to get one (having responded less than "College/university" on the question "What is the highest level of education you think you will actually get? (Don't know, No degree, Compulsory school, Upper secondary school, College/university)".

D   Figures

Figure D.1: Distribution of region of origin, LNU