EXPECTED EARLY LEAVING AMONG NATIVE AND MIGRANT STUDENTS: EVIDENCE FROM PISA FOR EU MEMBER STATES

Ralph Hippe and
Maciej Jakubowski

CEDEFOP

The **European Centre for the Development of Vocational Training** (Cedefop) is the European Union's reference centre for vocational education and training, skills and qualifications. We provide information, research, analyses and evidence on vocational education and training, skills and qualifications for policy-making in the EU Member States.
Cedefop was originally established in 1975 by Council Regulation (EEC) No 337/75. This decision was repealed in 2019 by Regulation (EU) 2019/128 establishing Cedefop as a Union Agency with a renewed mandate.

Europe 123, Thessaloniki (Pylea), GREECE
Postal: Cedefop service post, 570 01 Thermi, GREECE
Tel. +30 2310490111, Fax +30 2310490020
Email: info@cedefop.europa.eu
www.cedefop.europa.eu

Jürgen Siebel, *Executive Director*
Nadine Nerguisian, *Chair of the Management Board*

This working paper should not be reported as representing the views of the European Centre for the Development of Vocational Training (Cedefop). The views expressed are those of the authors and do not necessarily reflect those of Cedefop.

Acknowledgements

Contents

Tables

Tables

CHAPTER 1.
Introduction

Achieving equity and improving the education levels of young people is a policy priority for the European Union ([1]). The EU has set itself several targets in education: that at least 45% of 25-34 year-olds complete tertiary education, that at least 60% of vocational education and training graduates will have been exposed to work-based learning, and that the rates of early leaving from education and training are reduced to below 9% in the EU-27 by 2030 (Council of the European Union, 2021). Cedefop indicates that early leaving from education and training 'is understood at EU level as a failure to complete upper secondary education or a failure to gain qualifications or school leaving certificates' (Cedefop, 2016a, p. 1). It is a crucial issue as it 'is an obstacle to economic growth and employment. It hampers productivity and competitiveness, and fuels poverty and social exclusion' (European Commission, 2016, p. 1; see also Cedefop, 2016b). However, rates of early leavers from education and training (ELET) vary substantially among Member States (MS). According to the most recent Eurostat (2021) data, in 2020 Malta had the highest ELET rate (17%) in the EU, while Croatia had the lowest (2%) in 2020 ([2]).

Early leaving from education and training is a complex issue and a concern for policy-makers as empirical evidence also shows that leaving early is connected with several forms of individual and social disadvantage, with significant negative consequences for jobs, wages and life satisfaction (e.g. European Commission et al., 2014; Cedefop, 2016). In particular, immigrant students 'are largely over-represented among the early leavers from education and training in many European countries' (European Commission et al., 2014, p. 38; see also European Commission, 2021).

Persisting disparities between immigrants ([3]) and natives in education (and labour market) outcomes in MS are a source of inequality and social exclusion (see Flisi, Meroni, and Vera-Toscano, 2016; Rodrigues, 2018). Therefore, comparing

---

([1]) For more information on equity in education related issues, see Hippe; Araújo and Dinis da Costa (2016).

([2]) Note that ELET are defined as '18-24-year-olds with only lower secondary education or less who are no longer in education or training' (European Commission et al., 2014, p. 26). It can also "be distinguished from 'dropout', which refers to discontinuing an ongoing course, e.g. dropping out in the middle of the school term" (ibid.). The ELET definition thus differs from the PISA data we employ. See further details below.

([3]) In this paper, the terms immigrants and migrants are used synonymously.

educational expectations between immigrants and natives, at an early age, is particularly important if policy-makers intend to counteract disparities in educational achievement that later in life lead to social and economic inequalities.

The wave of large-scale refugee intake by MS since 2015 has increased the challenges associated with the integration of immigrants in the host countries (European Commission, 2016). The war in Ukraine has introduced a new dimension to this issue, with several million Ukrainians having had to flee their country within just a few days and weeks and seek refuge in MS following Russia's invasion. Many are children of school age (Psifidou and Treves, 2022). In these contexts, education is crucial to successful integration and employability (OECD, 2016b; Cedefop, 2021). At the same time, the free movement of people among MS has led to significant worker and student mobility within the EU. Our objective is to better understand inequalities related to (expected) early leaving from education and training within and across MS, giving special attention to migrant status.

To achieve this task, this paper uses the most recent available data from OECD's Programme for International Student Assessment (PISA) (⁴). As PISA 2015 and 2018 include a broader set of variables than previous rounds, we are able to consider a large range of possible relevant factors that may be related to early leaving from education and training (and migrant status). However, ELET rates cannot be measured directly in international student assessment data like PISA. Thus, it is only possible to consider students' expectations. Still, expected education levels provide useful information on potential future educational attainment. In fact, the relationship between expectations and (later) achievement is well established in the social sciences (Portes et al. 2010). Therefore, we can use these expectations to gain insights into the factors influencing early leaving from education and training.

While PISA 2015 and 2018 include data on students with immigrant background in all MS, for some MS the number of migrants in the sample is too low, so we cannot include them in our analysis. In addition, for a few MS there is no information on the educational expectations of students in PISA, which further reduces the sample. For these reasons, we analyse the expected education level in accordance with these sample size limitations.

We first analyse educational expectations by computing descriptive statistics and then by running a range of regression models for all MS together. More specifically, we use two-level logit regression models that include both the student

---

(⁴) PISA data are described in more detail in Section 3. For more information, see also the official OECD PISA website .

level and the school level to explore the factors that are significantly correlated with the likely expectation of leaving education and training early.

The results for all MS indicate that there are no inherent differences between natives and immigrants in Europe in the factors that influence the likelihood of expected early leaving from education and training. This finding implies that instead of focusing on immigrant-specific measures, it may be more appropriate to consider policy measures that deal with the underlying issues leading to the expectation of early leaving from education and training that are common to both natives and immigrant students. In fact, our regressions show that these underlying factors are related both to students and to schools. At the student level, the most relevant factors are the socioeconomic background of students, truancy and grade repetition. At the school level, it is not only the school's mean rate of expected early leavers from education and training but also other compositional effects related to student background.

The paper is divided as follows. First, we review the literature on immigrant student populations and educational expectations. The data and methodology are presented in the following section, which is followed by the empirical analysis. Finally, the conclusion sums up the results of the paper, provides policy recommendations and indicates future research avenues.

CHAPTER 2.
Literature review

2.1. Migrant students in Europe

The integration of immigrants across European education systems is a priority for policy-makers in the European Union (Hippe et al., 2016) ([5]). Similarly, the OECD (2016b, p. 243) stresses that '[m]igration puts enormous strains on both host communities and immigrants themselves; but it can also provide new opportunities for countries that face ageing native-born populations and the threat of labour and skill shortages'. In general, the educational achievement of immigrants lags behind that of native students in almost all MS. In most MS immigrant students come from less favourable backgrounds as their parents have, on average, lower levels of education than native individuals. In fact, evidence shows that socioeconomic status or family background are more strongly associated with students' performance than immigrant background (Blanden and McNally, 2015; Dustmann, Frattini and Lanzara, 2012; OECD, 2016b; Schnepf, 2007; Cedefop, 2016b). Clearly, it should be a priority for policy-makers to support and raise the achievement of low performing students, namely students with disadvantaged socioeconomic backgrounds, immigrant students and students from diverse ethnic minorities to reduce the impact of socioeconomic background on education outcomes.

Another factor that influences school performance is how long immigrant students have resided in the host country. Studies indicate that native students outperform immigrants, including those that arrived in the host country during their childhood (Dustmann and Theodoropoulos, 2010; Murat, 2011; OECD and European Union, 2015). In addition, there is also evidence that knowledge of the language of the host country, together with family background, are highly relevant as drivers of the immigrant-native gaps in many countries (Dustmann et al., 2012; OECD, 2016b). Specifically, language knowledge can help to reduce the immigrant-native educational gap.

On the influence of early tracking, Hanushek and Wößmann (2006) show that, in general, early tracking ([6]) of students in secondary schools, based on their academic skills, increases education inequality (see also Cedefop, 2016b). This is

---

([5]) Hippe et al. (2016) has been used for the following review of the literature on immigrant students.

([6]) Tracking in this context is the differentiation of school curricula into vocational and academic tracks.

in line with a study from Jakubowski and Pokropek (2015), which finds that while an early tracking policy might not be harmful for the best students, it can lower the performance of the weakest students, such as immigrants. There is also evidence that tracking at later stages does not always reduce score gaps between immigrant and native students, although it can contribute to improving the education opportunities of students who lack proficiency in the language of instruction (Ruhose and Schwerdt, 2016). In addition, in some EU MS early tracking systems are viewed as an obstacle to the integration of immigrants as, in general, the selection into tracks occurs before children become skilled in the language of the host country (De Paola and Brunello, 2016; Lüdemann and Schwerdt, 2010). Thus, Ruhose and Schwerdt (2016) suggest that action must be taken to improve the educational opportunities of children from less integrated families, calling for a more comprehensive school system that contributes to the integration of immigrant students.

The recent literature has also pointed out the importance of class and school composition in fostering immigrants' integration. Having a high share of immigrants in the class or school has a negative effect on immigrants' performance[7] (De Paola and Brunello, 2016). These results indicate that introducing a limit to the share of immigrant students in the classroom is a supportive policy that might reduce immigrants' education gap (similar policies have been introduced in Belgium and Denmark; see De Paola and Brunello, 2016) ([8]). Retaining and attracting more advantaged students in schools that also host immigrant students could be a supportive policy to improve immigrants' education opportunities (OECD, 2015b). In contrast, PISA 2015 results do not show a negative influence of immigrant concentration in schools on science scores, after controlling for the school's socioeconomic composition (OECD, 2016b).

The literature also reveals that free pre-school programmes for immigrants can help to increase equity for students with an immigrant background (De Paola and Brunello, 2016). Several recent system-level reforms have also been adopted to design more inclusive education systems, through structural changes to education systems or more targeted approaches, such as reducing grade repetition or raising the age of early tracking.

---

[7] Schneeweis (2015) shows that, for grade repetition and for tracking the attendance of students with an immigrant background, there is a negative association with a high share of immigrants, particularly for students of the same ethnic group.

[8] However, this may also lead to inequality and exclusion, so it needs to be assured that all immigrant students are given the chance of education and integration into mainstream schools.

In sum, the evidence shows that there are many factors influencing immigrants' integration and success at school and that, while some may be common across MS, others are country specific.

2.2.   Early leaving and educational expectations

Early leaving from education and training is an umbrella term which comprises several subcategories. These include: '(a) non-starters: those who decided not to continue or start any programme after completing education below the expected level (lower secondary education or short upper secondary education); (b) dropouts: those who started a programme that should lead them to the threshold qualification level but who dropped out before completion; (c) those who fail at final examinations, after completing the full programme' (Cedefop, 2016b, p. 2). Similarly, Cedefop has analysed early leavers and those students who are at risk of leaving early and found six distinct types: 'learners escaping the system', 'learners disengaging due to difficulties adapting after transition', 'learners who had to leave education and training because of caring, parenting or work obligations', 'learners confronting the system', 'learners disengaging because they cannot find a placement of their choice', and finally, 'learners combining multiple disadvantage, possibly facing health and psycho-social issues' (Cedefop, 2016a, p. 2; see also Cedefop, 2022c). This shows the diversity of situations which may produce early leavers. While many measures have been taken by EU MS for prevention, intervention, compensation, or a combination of measures with other policies (e.g. public employment services), and impact evaluation (for more details, see Cedefop, 2016c, d), more needs to be done. In particular, the evaluation of the impact of specific measures has only been done in a minority of policies and initiatives in the EU. Further, Cedefop recommends a strong policy commitment at both higher and lower levels to tackle early leaving and to make sure that any good practices are sustainable. Thus, cooperation at all levels, i.e. local, regional and national levels, is crucial to ensure that good practices are proliferated and effectively implemented (Cedefop, 2016a). To this end, Cedefop has also developed the VET toolkit for early leaving, providing resources for both policy-makers and practitioners to identify and monitor (potential) early leavers, intervene and evaluate the measures taken (Cedefop, 2022a) ([9]).

([9])   A similar toolkit has recently been released whose objective is to empower NEETs (those not in education, employment or training), see Cedefop (2022b). It is, therefore, complementary to the VET toolkit for tackling early leavers.

Early leaving is an important and policy-relevant issue, as it produces high costs for the individual and society (Brunello and De Paola, 2013; Cedefop, 2016b). ELET miss out on additional years in education and training, which is costly because 'an additional year of schooling can increase individual lifetime earnings by between 4% and 10%' (European Commission, 2013b, p. 11). Thus, individuals cannot reap the benefits of further education studies even though these advantages are manifold, not only in economic terms but also in other areas. For example, on average, individuals who stay longer in education have higher job satisfaction, take better informed decisions for their health and social life and increase their non-cognitive skills. The State also faces higher costs for its finances (e.g. potentially higher unemployment benefit payments) and its social welfare, as early leaving from education and training has detrimental impacts on crime rates and on attitudes towards immigrants and other minority groups (Brunello and De Paola, 2013). Therefore, better data on early leaving are needed at all levels to allow and improve targeted policy-making (Cedefop, 2016b).

The European Commission et al. (2014) names a number of factors that are most closely related to early leaving from education and training, such as low socioeconomic background, being an immigrant student and being male. In particular, early leaving from education and training is widespread among immigrant groups ([10]). Yet the most important determinant of early leaving from education and training is not immigrant status *per se*, but the often-associated low socioeconomic background.

However, in many international student assessments like PISA it is often not possible to measure the accomplished education level of students, only their expected level, as they are not followed through their subsequent education career once they have completed the assessment. That is, one cannot check whether expected education levels are identical with future realised education levels. In principle, there may be differences between the expected and the actually realised level, which may be influenced by personal ability, socioeconomic status (e.g. due to financial issues) and cultural values (which may be different between natives and (among different) migrant groups). Nevertheless, the literature emphasises

---

([10]) According to the European Commission et al. (2014, p. 39), '[t]he proportion of foreign-born compared to those born in the reporting country is nearly three to five times as high in Greece, Slovenia, Croatia, Austria and Switzerland. In Belgium, Spain, France, Italy, Cyprus, Finland and Sweden, the rates of early leavers among foreign-born students are still around twice as high. […] In Denmark, Ireland, Malta, the Netherlands and Portugal, the differences between the two groups are relatively lower, even though in some of these countries, high rates of migrants exist. Finally, one country stands out as an exception: the United Kingdom has a slightly higher rate of students leaving education and training early among those born in the country'.

that overall expected education levels correlate very highly with those actually realised. The analysis of expectations was first introduced about 50 years ago by Sewell and Shah (1968), Sewell; Haller and Ohlendorf (1970) and Sewell and Hauser (1972). These authors found in their models that parents' and students' educational expectations were predicting educational achievement quite well (see Gutierrez and Lopez-Agudo, 2016; Minello, 2014). This association has become established in the literature. In fact, authors such as Portes et al. (2010, p. 793) state that the 'relationship between expectations and achievement is arguably one of the best established facts in social science. The rationale is obvious: if a young person aims at some lofty goal, she may not achieve it; but if she does not aim high in the first place, she will surely not get there. Stated in this form, ambition becomes a prerequisite – a necessary condition – for achievement'. The OECD confirms these findings, indicating that '[e]xpectations for higher education and careers are often self-fulfilling prophecies: students who hold ambitious – but realistic – expectations for their future are more likely to put greater effort into their learning and make better use of the education opportunities available to them' (OECD, 2015a, p. 18).

At the same time, immigrant students may have differing education goals and expectations than their native counterparts. Their parents' decision to migrate was often motivated by the search for better work and education opportunities for themselves and their children (Goldenberg et al., 2001; Kao and Tienda, 1995; Phalet; Andriessen and Lens, 2004). While work opportunities may sometimes be more important than education prospects (Hagelskamp; Suárez-Orozco and Hughes, 2010; Massey and Taylor, 2004), a very good education for their children is often a key motivating factor for immigrant parents (Dustmann and Glitz, 2011; OECD, 2017a). Thus, the different backgrounds and the different reasons behind the decision to migrate may potentially induce immigrant students to have specific educational expectations. This is one of the reasons why this issue has recently received more attention by researchers (Feliciano, 2006; Minello, 2014; Minello and Barban, 2012). For example, the educational expectations of various immigrant groups tend to be high in a number of MS (Brinbaum and Cebolla-Boado, 2007; Jonsson and Rudolphi, 2010; Kristen and Dollmann, 2010; Relikowski; Yilmaz and Blossfeld, 2012; Salikutluk, 2016; Teney; Devleeshouwer and Hanquinet, 2013).

This finding can theoretically be explained in four different ways (see Salikutluk, 2016). The first is the immigrant optimism approach. Even if the parents were not successful and find themselves in the lower strata in the destination country, their children may continue the pursuit of their parents' goals, for which education is a key factor (Heath; Rothon and Kilpi, 2008; Tjaden and Hunkler,

2017). Second, immigrants may react positively to difficulties in the destination country by trying to overcome these hurdles, as summarised in the blocked opportunities assumption (e.g. Heath and Brinbaum, 2007). Third, it is possible that immigrant parents do not have sufficient information about the education system of the country they have migrated to, so they underestimate the requirements and standards for high educational attainment (Kao and Tienda, 1995). Finally, the so-called influence of significant others hypothesis states that close relationships within and outside the family (parents, siblings, peers, friends,) may have various positive and negative influences on the expectations of students, and immigrant students in particular (e.g. Gabay-Egozi; Shavit and Yaish, 2014; Nauck and Kohlmann, 1999).

In addition, many other factors may be related to educational expectations. For example, the OECD notes that '[s]tudents' expectations of further education are influenced by education policy, particularly the degree of sorting students into different education tracks' (OECD, 2017a, p. 104). Thus, in our following analyses we will also assess various ways in which education policy is related to educational expectations.

CHAPTER 3.
Data and methodology

3.1.   Variables included in this study

For the purposes of this paper, we have included a set of variables that have been used previously when analysing the performance of students with an immigrant background and students' expectations of further education (OECD, 2015a; 2016b; 2017a). In addition, we have included one variable (truancy index) which is not available in the PISA data sets, and which was estimated for the purposes of this research, as it is highly relevant for showing the achievement of low-performing students (Hippe; Jakubowski and Araújo, 2018).

The key outcome variable related to expected early leaving from education and training is measured via the student responses to the question: Which of the following do you expect to complete? (see also Annex 2), which allows students to choose from different ISCED levels. However, while the question and response categories were exactly the same in PISA 2015 and 2018, in 2015 students needed to select one response, while in 2018 they could select all responses and each response was separately available in the data. For our purposes, we code expected early leaving from education and training in a binary way, which means that leaving early is either expected (when students choose the ISCED 2 category) or not expected (in those cases where students choose a level higher than ISCED 2). This cut-off point has been chosen according to the official EU definition of early leavers from education and training who have ISCED 2 as their highest school attainment level. For PISA 2018 data, we took the highest ISCED level clicked by a student. Thus, if a student clicked ISCED 2 only, then we coded her/him as 1 for the outcome variable. However, if in PISA 2018 this student also clicked ISCED 3, we coded it 0 for the outcome variable.

It needs to be stressed that there is a difference between official ELET rates, as published by Eurostat, and the OECD PISA data used in this study. This means that the outcome variable is not exactly the concept of early leaving as defined for EU policies and statistics: only those attending school are included in the PISA sample, the age groups are different (18-24-year-olds in Eurostat while PISA includes the expectations of 15-year-olds), participation in non-formal training is included in the EU concept but it is not within the scope of the PISA data analysis. Thus, results from this study cannot be put one to one on ELET rates in the Eurostat sense. Still, the PISA data allow us to better understand expectations

about early leaving, and this is insightful as there is a relationship between expectations and (later) achievement (Portes et al., 2010).

Table 1 provides a brief description of all variables. To simplify the reading of the regression tables, we also include information on the reference category in the multilevel regressions. For continuous indices, we also provide information on how to interpret higher values on this index.

Table 1.   **Description of student-level variables**

Variable	Description
Expected early leaving from education and training	Leaving school early (after ISCED 2) is expected or not. Reference category: not expecting to leave school early.
Female	Female=1, Male=0.
Student's ESCS	Measures student socioeconomic background. Positive values are reflecting higher socioeconomic status of a student family.
Migrant	Denotes students with an immigrant background. Reference category: native.
1st generation	1st generation immigrant or native student. Reference category: native.
2nd generation	2nd generation immigrant or native student. Reference category: native.
Language at home	Language taught at school and the language spoken at home are different. Reference category: languages are not different.
Durecec	Duration in early childhood education and care (in years).
Sense of belonging	Measures how well students feel integrated and belonging to the school they attend (higher values for students with high sense of belonging).
Truancy	Measures whether students are skipping school days, skipping classes or being late at school (higher values for students who often skip classes or are late at school).
Grade repetition	Student has repeated a grade or not. Reference category: not having repeated a grade.
Student achievement	Achievement in science (for 2015) or reading (for 2018) (1st plausible value).
(Programme) Designation	Student attends: general programmes designed to give access to the next programme level, or programmes designed to give access to vocational studies at the next programme level, or giving direct access to labour market, or modular programmes. Reference category: general programmes.
(Programme) Orientation	Student attends: general, pre-vocational or vocational programmes (and in Luxembourg, also modular programmes). Reference category: general.
ISCED Level	Student attends: ISCED 2 or ISCED 3. Reference category: ISCED 2.

Table 2 provides descriptive statistics and sample size variables for the variables used in the analysis. Due to missing questionnaire responses for some students, as well as for entire countries for some variables, the sample size is lower than that available for the estimation of achievement scores (full PISA sample). Depending on the regression model, we estimate results for all students with data available for this particular set of predictors. For example, Spain does not have achievement data for 2018, so Spanish students are not included in the regression with achievement data, but they are included in other estimations. Overall, the data are available for more than 260 000 students from more than 10 000 schools in Europe. The final estimation sample sizes are provided in regression tables.

Means and standard deviations are provided separately for 2015 and 2018 and also for the combined data. The expected early leaving from education and training indicator is distributed differently across 2015 and 2018. This might be due to the different way of collecting responses in these two studies (see above) but might also reflect changes in student expectations. In 2015, the average is higher, suggesting that 13% of all students expect to finish their education early. In 2018 that share goes down to 10%. Also, the share of students with an immigrant background increased from 12% to 14%, an increase mainly due to a larger share of second-generation immigrant students. For other variables, distributions are similar across 2015 and 2018, suggesting that the analysis based on the combined data set is plausible.

Table 2.   **Descriptive statistics for PISA 2015 and 2018 variables**

Variable	2015			2018			2015 and 2018		
	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.	N
Expected early leaving	0.13	0.34	132 379	0.10	0.30	157 574	0.11	0.32	289 953
Gender	0.50	0.50	137 120	0.49	0.50	164 146	0.49	0.50	301 266
Student's ESCS	-0.06	0.96	134 050	-0.05	0.96	160 166	-0.06	0.96	294 216
Migrant	0.12	0.33	133 042	0.14	0.35	159 187	0.13	0.34	292 229
1st generation	0.05	0.22	133 042	0.05	0.22	159 187	0.05	0.22	292 229
2nd generation	0.07	0.26	133 042	0.09	0.29	159 187	0.08	0.27	292 229
Language at home	0.12	0.32	134 697	0.14	0.35	161 166	0.13	0.34	295 863
Durecec	2.92	1.04	105 724	2.99	1.00	129 785	2.96	1.02	235 509
Sense of belonging	0.13	0.99	132 284	0.11	0.97	149 230	0.12	0.98	281 514
Truancy	-0.03	0.97	130 429	0.12	1.11	133 027	0.04	1.04	263 456
Grade repetition	0.19	0.39	133 410	0.17	0.37	158 993	0.18	0.38	292 403

Variable	2015			2018			2015 and 2018		
	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.	N
Student achievement	495.9	97.3	137 120	490.4	101.7	128 203	493.3	99.5	265 323
ISCED Level	2.47	0.50	137 119	2.49	0.50	157 344	2.48	0.50	294 463
(Programme) Designation	1.17	0.47	137 119	1.22	0.49	157 344	1.19	0.48	294 463
(Programme) Orientation	1.36	0.75	137 119	1.34	0.73	157 344	1.35	0.74	294 463

3.2. Students with a migrant background and the final estimation sample

In many EU MS, students with an immigrant background constitute a substantial share of all 15-year-olds, but in some the number of such students is relatively low and the sample size available in PISA is too small to provide reliable inferences. We excluded data from Bulgaria, Poland, Romania, and Slovakia, as these MS had less than 100 students with an immigrant background in 2015 or 2018, or in both years. Also, Malta did not collect responses about expected education level in 2015 and the data for Cyprus are not included in the public databases, so these two MS are not included in the analysis. Except for these six MS, our data set includes results for all other EU MS from both PISA 2015 and 2018.

Table 3 provides sample sizes with all students and those with an immigrant background, separately for each country included in the final estimation sample, by PISA wave, and also for the combined 2015 and 2018 sample ([11]). The final results are weighted by survey probability weights, so they are representative of the underlying populations of 15-year-old students in the MS listed in Table 3. As only six EU MS were excluded from the analysis ([12]) and relative to the whole EU population none of them has a substantial number of students with an immigrant background, one could argue that this sample and the results are representative of the population of students with an immigrant background in the European Union.

([11]) This was done to see the overall results as well as for each year separately.

([12]) Due to an insufficient number of students with an immigrant background in these countries.

Table 3. **Sample size and the number of students with an immigrant country by country and PISA wave.**

MS	All students			Students with an immigrant background		
	2015	2018	2015 and 2018	2015	2018	2015 and 2018
Austria	7 007	6 802	13 809	1 319	1 407	2 726
Belgium	9 651	8 475	18 126	1 551	1 473	3 024
Croatia	5 809	6 609	12 418	605	598	1 203
Czechia	6 894	7 019	13 913	226	251	477
Denmark	7 161	7 657	14 818	1 686	1 551	3 237
Estonia	5 587	5 316	10 903	560	539	1 099
Finland	5 882	5 649	11 531	230	313	543
France	6 108	6 308	12 416	762	947	1 709
Germany	6 504	5 451	11 955	967	1 050	2 017
Greece	5 532	6 403	11 935	527	702	1 229
Hungary	5 658	5 132	10 790	147	124	271
Ireland	5 741	5 577	11 318	760	958	1 718
Italy	11 583	11 785	23 368	899	1 071	1 970
Latvia	4 869	5 303	10 172	243	244	487
Lithuania	6 525	6 885	13 410	221	148	369
Luxembourg	5 299	5 230	10 529	2 672	2798	5 470
Netherlands	5 385	4 765	10 150	560	661	1 221
Portugal	7 325	5 932	13 257	422	342	764
Slovenia	6 406	6 401	12 807	518	573	1 091
Spain	6 736	35 943	42 679	681	4 117	4 798
Sweden	5 458	5 504	10 962	898	1 055	1 953

3.3. Statistical model

Multilevel regression models are well-suited to the analysis of large-scale international student assessment surveys because they recognise the hierarchical structure of the data with students nested in schools. Accordingly, in our analysis the two-level model reflects that students are nested in schools. In addition, in most regressions we control for country fixed effects to exclude all student- and school-invariant characteristics shared by a student population in each country. This seems to be important in the analysis of expected early leaving from education and training as countries' fixed school structure and labour market regulation affect the

choices of students and are not related to student- or school-level factors which we associate with early leaving decisions.

The basic two-level model with random effects can be described by two equations. The first equation describes the model for the student-level where i is an index for students and j is an index for schools:

$$y_{ij} = \beta_{0j} + \beta_1 \mathrm{x}_{ij} + e_{ij}.$$

In this equation, y_{ij} is the outcome variable, x_{ij} is a vector of student-level characteristics and e_{ij} is a student-level error. The second level equation describes school level intercepts:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} W_j + u_{0j},$$

where w_j is a set of school-level predictors and u_{0j} is a random error component at the school-level.

In our case the outcome variable is the expected early leaving from education and training indicator, which takes the value 0 (i.e. early leaving from education and training is not expected) or 1 (i.e. early leaving from education and training is expected). We apply a multilevel binary logistic model (or logistic regression with random intercepts), which, by using the above notation, can be defined in one equation as:

$$\mathrm{logit}\{\mathrm{Pr}\,(y_{ij} = 1 | \mathrm{x}_{ij}, u_{0j})\} = \gamma_{00} + \gamma_{01} \mathrm{w}_j + \beta_1 \mathrm{x}_{ij} + u_{0j}.$$

We assume that the school-specific random term is independently distributed across schools following $u_{0j} | \mathrm{x}_{ij} \sim N(0, \varphi)$ and that y_{ij} are independently distributed following conditional binomial distribution.

The associations between performance and student-level or school-level factors can also be decomposed into within-school and between-school associations. Thus, with these models it is possible to analyse how differences in key policy-relevant variables are associated with student and school characteristics. In our case, we centre student-level predictors at weighted school means and we include school averages in the regressions. All models are weighted with student-level probability weights but also with school-level probability weights, to reflect different sampling probabilities at the two stages of complex sampling in the PISA survey (see Rabe-Hesketh and Skrondal, 2006, for further discussion of survey weighting with PISA data).

CHAPTER 4.
Empirical analysis

	2015	2018	2015 and 2018	2015	2018	2015 and 2018
	(1)	(2)	(3)	(4)	(5)	(6)
First generation	-0.01	0.25	0.03			
	(0.16)	(0.39)	(0.15)			
Second generation	-0.17	0.16	-0.07			
	(0.18)	(0.30)	(0.16)			
Speaking other language at home	0.26*	0.15	0.24*			
	(0.12)	(0.26)	(0.12)			
2018.wave dummy			-0.65***			-0.64***
			(0.16)			(0.16)
Female				0.02	-0.26	-0.05
				(0.11)	(0.14)	(0.09)
ESCS				-0.57***	-0.55***	-0.55***
				(0.04)	(0.09)	(0.04)
Migrant				-0.21	-0.03	-0.17
				(0.13)	(0.21)	(0.11)
ISCED 2 (baseline category)						
ISCED 3				-3.20***	-1.57***	-2.67***
				(0.21)	(0.39)	(0.21)
ISCED designation = general (baseline category)						
ISCED designation = further vocational education				0.31	0.91	0.85**
				(0.52)	(0.58)	(0.32)
ISCED designation = labour market				0.72	1.14	0.98**
				(0.60)	(0.61)	(0.36)
ISCED orientation = general (base category)						
ISCED orientation = pre-vocational				-0.88	0.64	-0.64
				(0.55)	(0.55)	(0.37)
ISCED orientation = vocational				0.73*	1.62*	1.15***
				(0.37)	(0.67)	(0.32)
ISCED orientation = modular				1.08***	1.12***	1.04***
				(0.28)	(0.29)	(0.20)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-3.85***	-4.60***	-4.00***	-1.70***	-4.11***	-2.59***
	(0.42)	(0.27)	(0.44)	(0.33)	(0.45)	(0.37)
Variance at the school level	1.94	3.67	2.44	1.13	2.60**	1.61***
	(0.18)	(0.96)	(0.30)	(0.13)	(0.90)	(0.26)
N of students	99823	77313	177136	99823	77313	177136

NB: Standard errors in parentheses; * p<0.05, ** p<0.01, *** p<0.001

Table 6 provides regression results testing whether the association between expected early leaving from education and training and socioeconomic

background varies between native and immigrant students. There are six regressions estimated on different samples and with and without dummy indicators for school level, programme orientation and designation. In all regressions there is no difference in the expectation of early leaving from education and training between native and migrant students, and the socioeconomic background is negatively related to this expectation. In all regressions, the interaction term between ESCS and a dummy indicating students with an immigrant background is positive, although significant for both specifications only in the pooled 2015 and 2018 sample. This suggests that the relationship between socioeconomic background and expected early leaving from education and training is weaker for immigrant students. Thus, contrary to many beliefs, this seems to suggest that based on our data immigrant students' educational expectations are less affected by their family education, economic, and cultural capital, at least when it comes to decisions related to whether to continue education or not.

Table 6. **Models estimating the difference in the ESCS slope between native and migrant students**

	PISA 2015		PISA 2018		2015 and 2018	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.01	0.02	-0.30*	-0.26	-0.08	-0.05
	(0.11)	(0.11)	(0.13)	(0.14)	(0.09)	(0.09)
ESCS	-0.63***	-0.61***	-0.66***	-0.59***	-0.64***	-0.59***
	(0.05)	(0.05)	(0.11)	(0.11)	(0.05)	(0.04)
Migrant	-0.03	-0.06	0.09	0.09	-0.01	-0.04
	(0.14)	(0.14)	(0.20)	(0.21)	(0.12)	(0.12)
Migrant*ESCS	0.24*	0.23	0.21	0.17	0.22*	0.19*
	(0.12)	(0.12)	(0.16)	(0.16)	(0.09)	(0.09)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
PISA 2018 data dummy					Yes	Yes
ISCED level fixed effect		Yes		Yes		Yes
ISCED orientation fixed effect		Yes		Yes		Yes
ISCED designation fixed effect		Yes		Yes		Yes
Constant	-3.71***	-1.69***	-4.30***	-4.13***	-3.80***	-2.59***
Variance at the school level	1.48***	1.12***	2.95**	2.60**	1.88***	1.61***
N of students	99 823	99 823	77 313	77 313	177 136	177 136

NB: Standard errors in parentheses; * p<0.05, ** p<0.01, *** p<0.001

Table 7 provides results for a fuller multilevel model with a set of student- and school-level indicators. In all regressions, the duration of pre-primary education and care is not related to early leaving from education and training decisions. That might be counterintuitive as it is believed that preschool education is important, especially for disadvantaged students who, as we see, consider finishing

education early more often. However, these are data reported by 15-year-olds for whom the effects of pre-primary education may be covered by other more relevant factors at this age.

Also, the index of school belonging seems not to be associated with early leaving from education and training decisions. Only in regression (4) is this index statistically significant, but with a relatively large standard error and a positive coefficient, which is not what one could expect. In all other regressions the association is insignificant despite a large sample, suggesting that relations at school may not matter that much for student decisions about their education career.

Truancy, or skipping classes and being late to school, appears to be a good predictor of expected early leaving from education and training. This is not a surprising finding, but it is of practical importance. Students who show similar behaviour should be supported by schools, if possible, to avoid leaving education and training altogether.

Grade repetition is also found to be associated with the likelihood of higher expected early leaving, which can be explained by the fact that students repeating a grade are those who in fact struggle with learning more than their peers, and thus are more susceptible to early school leaving. In some EU MS grade repetition is rare, despite the fact that these MS might face similar challenges for students' low performance or socioeconomic background. This makes us think that there might be room for rethinking grade repetition practices in MS.

In these multilevel models, we separate the effects of socioeconomic background (ESCS) into within-school individual student effect and school composite effect. Both are negatively related to expected early leaving, but while the individual effect remains similar across samples and specifications, the between-school composite effect is much weaker after including other school-average composite indicators: school average achievement and school average likelihood of expected early leaving from education and training. The latter is consistently positively related to expected early leaving from education and training across different samples. Thus, while the compositional effect of student ESCS is associated with early leaving from education and training decisions, it is partly related to the compositional effects of the popularity of expected early leaving in a school. Relative to other schools with similar ESCS background, the likelihood of expected early leaving from education and training for an individual student increases if other students in the same school consider finishing school early. Obviously, this association is driven by school or student body characteristics that cannot be disentangled in large-scale surveys, but the strength of this relationship

suggests that policies supporting schools with a large number of leavers are necessary to encourage individual students to continue education.

In contrast, the association with student achievement seems to be complex. At the individual level, higher achievement lowers the likelihood of early leaving. This is intuitive as students who do well are more likely to continue education. However, the compositional effect at the school level is positive, which means that in schools of similar socioeconomic background, with a similar share of students considering early leaving, and controlling for other characteristics in our model, higher average achievement appears to increase the likelihood of expected early leaving for an individual student. One possible explanation is that in higher performing schools some students may be thinking about leaving education early as they perceive their own performance negatively, a phenomenon that could be similar to the big-fish-little-pond effect (Fang et al., 2018).

The multilevel regressions also include an individual level dummy for immigrant students and school-average share of immigrant students. Both are insignificant or negatively related to the possibility of early leaving after controlling for other compositional and individual effects, and also for school characteristics. Based on the available data, these results suggest that – after controlling for individual and compositional effects of other variables – we do not find evidence of a significant positive relationship between having a migrant background and declaring an intention to abandon formal education early, even in schools with more students having an immigrant background.

Finally, note that the models in columns (2), (4) and (6) explain almost all variance at the school level. Thus, the compositional effects specified in our regressions are almost entirely explaining variation in early leaving from education and training decisions across schools.

Table 8 provides additional results comparing associations between native and migrant students. Grade repetition and truancy have similar associations with native and immigrant students. In a fuller specification in columns (4) to (6), within- and between-school associations of socioeconomic background and expected early leaving from education and training are also similar across native and immigrant students. The peer effects of having high achieving students and colleagues who consider early leaving from education and training are also weaker for migrant students. The latter association is significantly lower in all three regressions for immigrant students. Overall, these additional results demonstrate small differences between native and migrant students, and when significant, compositional and individual effects that are positively associated with expected early leaving from education and training tend to be smaller for immigrant students.

Table 7. **Multilevel models with additional student and composite predictors**

	PISA 2015		PISA 2018		2015 and 2018	
	(1)	(2)	(3)	(4)	(5)	(6)
Duration of early education and care	-0.00 (0.04)	0.00 (0.04)	0.13 (0.08)	0.07 (0.07)	0.03 (0.04)	0.01 (0.03)
Sense of belonging	-0.00 (0.05)	0.01 (0.05)	0.09 (0.07)	0.14* (0.07)	0.02 (0.04)	0.05 (0.04)
Truancy	0.12** (0.04)	0.08 (0.04)	0.26*** (0.05)	0.16*** (0.05)	0.18*** (0.03)	0.11*** (0.03)
Grade repetition	1.04*** (0.29)	0.56* (0.26)	0.36* (0.17)	0.16 (0.16)	0.81*** (0.21)	0.42* (0.17)
Female	0.06 (0.11)	-0.00 (0.11)	-0.22 (0.14)	-0.20 (0.13)	-0.00 (0.09)	-0.03 (0.10)
Within-school ESCS effect	-0.40*** (0.05)	-0.36*** (0.05)	-0.46*** (0.08)	-0.44*** (0.09)	-0.42*** (0.04)	-0.38*** (0.05)
Between-school ESCS effect	-1.90*** (0.15)	-0.15* (0.07)	-2.68*** (0.39)	-0.54* (0.26)	-2.11*** (0.18)	-0.28* (0.11)
Migrant	-0.23 (0.14)	-0.49** (0.15)	0.07 (0.21)	0.03 (0.22)	-0.14 (0.12)	-0.30* (0.13)
School share of migrant students	-0.08 (0.29)	-0.10 (0.18)	-2.59** (0.84)	-1.07* (0.45)	-1.06** (0.37)	-0.40 (0.21)
Student achievement (science or reading)		-0.01*** (0.00)		-0.00*** (0.00)		-0.01*** (0.00)
School average achievement		0.01*** (0.00)		0.00* (0.00)		0.01*** (0.00)
School share of students expecting early leaving		5.95*** (0.17)		6.90*** (0.76)		6.16*** (0.26)
2018 wave dummy					Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-4.35*** (0.34)	-3.02*** (0.41)	-4.40*** (0.36)	-3.94*** (1.12)	-4.12*** (0.33)	-3.25*** (0.53)
Variance at the school level	0.74***	0.00*	1.74***	0.25*	1.12***	0.00
N of students	99 823	99 823	77 313	77 313	177 136	177 136

NB: Standard errors in parentheses; * p<0.05, ** p<0.01, *** p<0.001

Table 8.   **Additional results with interaction terms for migrant students**

	2015	2018	2015 and 2018	2015	2018	2015 and 2018
	(1)	(2)	(3)	(4)	(5)	(6)
Truancy	0.12**	0.26***	0.18***	0.08	0.13*	0.09**
	(0.04)	(0.05)	(0.03)	(0.05)	(0.05)	(0.03)
Grade repetition	1.04***	0.35*	0.81***	0.45	0.24	0.37
	(0.30)	(0.17)	(0.22)	(0.30)	(0.19)	(0.20)
Female	0.06	-0.21	-0.00	-0.00	-0.20	-0.03
	(0.11)	(0.13)	(0.09)	(0.11)	(0.13)	(0.10)
Within-school ESCS effect	-0.44***	-0.48***	-0.46***	-0.39***	-0.46***	-0.41***
	(0.05)	(0.10)	(0.05)	(0.05)	(0.11)	(0.05)
Between-school ESCS effect	-1.86***	-2.66***	-2.08***	-0.15*	-0.52*	-0.28*
	(0.15)	(0.40)	(0.18)	(0.08)	(0.26)	(0.12)
Migrant	-0.32*	0.04	-0.20	-0.79***	0.07	-0.44*
	(0.15)	(0.25)	(0.13)	(0.22)	(0.25)	(0.17)
School share of migrant students	-0.21	-2.64**	-1.15**	-0.41*	-0.96*	-0.55*
	(0.31)	(0.88)	(0.38)	(0.20)	(0.49)	(0.22)
Migrant*ESCS	0.23	0.09	0.18	0.16	0.10	0.12
	(0.13)	(0.16)	(0.10)	(0.14)	(0.18)	(0.11)
Migrant*school average ESCS	-0.54**	-0.15	-0.39*	-0.20	-0.12	-0.16
	(0.20)	(0.36)	(0.18)	(0.25)	(0.35)	(0.21)
Migrant*truancy				0.03	0.16	0.09
				(0.11)	(0.13)	(0.08)
Migrant*repeat				0.66	-0.34	0.26
				(0.37)	(0.41)	(0.27)
Within-school achievement				-0.01***	-0.00***	-0.01***
				(0.00)	(0.00)	(0.00)
School average achievement				0.01***	0.00*	0.01***
				(0.00)	(0.00)	(0.00)
Migrant*school average achievement				-0.00	-0.01*	-0.00*
				(0.00)	(0.00)	(0.00)
School average share of expected early leavers from education and training				5.98***	6.93***	6.18***
				(0.17)	(0.78)	(0.26)
Migrant*school share of leavers				-1.56***	-2.15*	-1.72***
				(0.46)	(1.00)	(0.45)
2018 wave dummy			Yes			Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-4.39***	-3.94***	-4.05***	-3.07***	-3.83***	-3.26***
	(0.37)	(0.25)	(0.33)	(0.39)	(1.08)	(0.51)
Variance at the school level	0.73***	1.74***	1.12***	0.00**	0.26*	0.00
N of students	99823	77313	177136	99823	77313	177136

NB: Standard errors in parentheses; * p<0.05, ** p<0.01, *** p<0.001

shows the results from the two-level logit regression with school random effects
and students being nested in schools, without student- or school-level predictors.
This is the so-called empty model that provides a basis for comparisons with
predictors added. Models (1), (3) and (5) do not include any explanatory variables
except a dummy controlling for differences between 2015 and 2018 in the pooled
sample. These empty models show a substantial variation at the school level,
which confirms a multilevel approach. By comparing to models (2), (4) and (6),
which differ only by inclusion of the country fixed effects, it seems that around 40%
of the school-level variance is explained by fixed between-country differences. As
we are not interested in fixed between-country characteristics, but rather in
individual and school-level factors affecting expected early leaving from education
and training, we will include country fixed effects in all remaining regressions.

Table 4.   **Reference multilevel models without student- or school-level**
           predictors

	PISA 2015		PISA 2018		2015 and 2018	
	(1)	(2)	(3)	(4)	(5)	(6)
Dummy for the 2018 data					-0.61***	-0.65***
					(0.16)	(0.16)
Country fixed effects	No	Yes	No	Yes	No	Yes
Constant	-2.16***	-3.81***	-3.11***	-4.53***	-2.24***	-3.95***
	(0.08)	(0.42)	(0.16)	(0.27)	(0.10)	(0.44)
Variance at the school level	3.31***	1.95***	5.73***	3.69***	4.02***	2.46***
	(0.29)	(0.18)	(1.51)	(0.95)	(0.47)	(0.30)
N of students	99 823	99 823	77 313	77 313	177 136	177 136
N of schools	4 827	4 827	4 765	4 765	9 592	9 592

NB: Standard errors in parentheses; * p<0.05, ** p<0.01, *** p<0.001

Table 5 shows multilevel regressions with student background and school
type variables. The first three regressions show results with three separate
indicators related to immigrant background. Both dummies denoting first- and
second-generation students are insignificant after controlling for students who
speak a different language at home than the language of the PISA assessment.
This last dummy indicator is positively associated with the expected early leaving
from education and training but results are not significant for 2018 data and
relatively large standard errors suggest a large variation among students.

Columns (4)-(6) provide results for our preferred specification with only one
dummy variable indicating immigrant background (for both first- and second-
generation students) and with gender and socioeconomic status (SES) controlled

for. Only the ESCS index, which is a PISA index to measure SES, is significantly related to the expected early leaving from education and training, which confirms our intuition that students coming from more affluent families have, in general, a lower likelihood of finishing education early. The results suggest immigrants may not be special or inherently different in their expected early leaving behaviour relative to natives. Both immigrants and natives behave in the same way when forming their educational expectations, influenced by several common factors that we control for.

Regressions in columns (4)-(6) also control for school ISCED level, designation and orientation ([13]). As expected, if a student is already at ISCED 3, the odds of early leaving (i.e. finishing after ISCED 2) are lower than for ISCED 2 students. This result is intuitive, as such a student is already attending a higher school level. Attending a school that has a vocational or partly vocational programme orientation or designation increases the likelihood of early school leaving. Our results suggest that students who attend vocational programmes are more likely to consider early school leaving before finishing ISCED 3 ([14]).

Table 5.   **Multilevel models with student background and school characteristics**

	2015	2018	2015 and 2018	2015	2018	2015 and 2018
	(1)	(2)	(3)	(4)	(5)	(6)
First generation	-0.01 (0.16)	0.25 (0.39)	0.03 (0.15)			
Second generation	-0.17 (0.18)	0.16 (0.30)	-0.07 (0.16)			
Speaking other language at home	0.26* (0.12)	0.15 (0.26)	0.24* (0.12)			

([13]) To Cedefop's knowledge, PISA is not explicitly designed to ensure reliable comparable results for levels and orientation of education, including upper secondary VET. The use of the orientation variable has also been strongly debated in the past. Thus, there are caveats that have to be taken into account; the results need to be interpreted with caution.

([14]) This finding may be further contextualised: as Cedefop (2016b) shows, vocational education and training (VET) suffers from high rates of early leaving (in the Eurostat sense) in the first 3 to 6 months of students starting their VET programme. This is often due to a lack of career guidance for choosing the most appropriate pathway for them. Students who were at risk of dropping out from general education had a higher chance of completing their programme and qualifying in upper secondary education when they changed to VET. In consequence, VET may prevent early leaving. Students may also want to do a VET programme explicitly to start working directly after ISCED 2. Cedefop (2016b) also provides evidence that those who drop out and wish to go back to education choose VET and subsequently qualify. This means that VET also compensates the phenomenon.

	2015	2018	2015 and 2018	2015	2018	2015 and 2018
	(1)	(2)	(3)	(4)	(5)	(6)
2018.wave dummy			-0.65***			-0.64***
			(0.16)			(0.16)
Female				0.02	-0.26	-0.05
				(0.11)	(0.14)	(0.09)
ESCS				-0.57***	-0.55***	-0.55***
				(0.04)	(0.09)	(0.04)
Migrant				-0.21	-0.03	-0.17
				(0.13)	(0.21)	(0.11)
ISCED 2 (baseline category)						
ISCED 3				-3.20***	-1.57***	-2.67***
				(0.21)	(0.39)	(0.21)
ISCED designation = general (baseline category)						
ISCED designation = further vocational education				0.31	0.91	0.85**
				(0.52)	(0.58)	(0.32)
ISCED designation = labour market				0.72	1.14	0.98**
				(0.60)	(0.61)	(0.36)
ISCED orientation = general (base category)						
ISCED orientation = pre-vocational				-0.88	0.64	-0.64
				(0.55)	(0.55)	(0.37)
ISCED orientation = vocational				0.73*	1.62*	1.15***
				(0.37)	(0.67)	(0.32)
ISCED orientation = modular				1.08***	1.12***	1.04***
				(0.28)	(0.29)	(0.20)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-3.85***	-4.60***	-4.00***	-1.70***	-4.11***	-2.59***
	(0.42)	(0.27)	(0.44)	(0.33)	(0.45)	(0.37)
Variance at the school level	1.94	3.67	2.44	1.13	2.60**	1.61***
	(0.18)	(0.96)	(0.30)	(0.13)	(0.90)	(0.26)
N of students	99823	77313	177136	99823	77313	177136

NB: Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6 provides regression results testing whether the association between expected early leaving from education and training and socioeconomic background varies between native and immigrant students. There are six regressions estimated on different samples and with and without dummy indicators for school level, programme orientation and designation. In all regressions there is no difference in the expectation of early leaving from education and training between native and migrant students, and the socioeconomic background is negatively related to this expectation. In all regressions, the interaction term between ESCS and a dummy indicating students with an immigrant background is positive, although significant for both specifications only in the pooled 2015 and 2018 sample. This suggests that the relationship between socioeconomic

background and expected early leaving from education and training is weaker for immigrant students. Thus, contrary to many beliefs, this seems to suggest that based on our data immigrant students' educational expectations are less affected by their family education, economic, and cultural capital, at least when it comes to decisions related to whether to continue education or not.

Table 6. **Models estimating the difference in the ESCS slope between native and migrant students**

	PISA 2015		PISA 2018		2015 and 2018	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.01	0.02	-0.30*	-0.26	-0.08	-0.05
	(0.11)	(0.11)	(0.13)	(0.14)	(0.09)	(0.09)
ESCS	-0.63***	-0.61***	-0.66***	-0.59***	-0.64***	-0.59***
	(0.05)	(0.05)	(0.11)	(0.11)	(0.05)	(0.04)
Migrant	-0.03	-0.06	0.09	0.09	-0.01	-0.04
	(0.14)	(0.14)	(0.20)	(0.21)	(0.12)	(0.12)
Migrant*ESCS	0.24*	0.23	0.21	0.17	0.22*	0.19*
	(0.12)	(0.12)	(0.16)	(0.16)	(0.09)	(0.09)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
PISA 2018 data dummy					Yes	Yes
ISCED level fixed effect		Yes		Yes		Yes
ISCED orientation fixed effect		Yes		Yes		Yes
ISCED designation fixed effect		Yes		Yes		Yes
Constant	-3.71***	-1.69***	-4.30***	-4.13***	-3.80***	-2.59***
Variance at the school level	1.48***	1.12***	2.95**	2.60**	1.88***	1.61***
N of students	99 823	99 823	77 313	77 313	177 136	177 136

NB: Standard errors in parentheses; * p<0.05, ** p<0.01, *** p<0.001

Table 7 provides results for a fuller multilevel model with a set of student- and school-level indicators. In all regressions, the duration of pre-primary education and care is not related to early leaving from education and training decisions. That might be counterintuitive as it is believed that preschool education is important, especially for disadvantaged students who, as we see, consider finishing education early more often. However, these are data reported by 15-year-olds for whom the effects of pre-primary education may be covered by other more relevant factors at this age.

Also, the index of school belonging seems not to be associated with early leaving from education and training decisions. Only in regression (4) is this index statistically significant, but with a relatively large standard error and a positive coefficient, which is not what one could expect. In all other regressions the association is insignificant despite a large sample, suggesting that relations at

school may not matter that much for student decisions about their education career.

Truancy, or skipping classes and being late to school, appears to be a good predictor of expected early leaving from education and training. This is not a surprising finding, but it is of practical importance. Students who show similar behaviour should be supported by schools, if possible, to avoid leaving education and training altogether.

Grade repetition is also found to be associated with the likelihood of higher expected early leaving, which can be explained by the fact that students repeating a grade are those who in fact struggle with learning more than their peers, and thus are more susceptible to early school leaving. In some EU MS grade repetition is rare, despite the fact that these MS might face similar challenges for students' low performance or socioeconomic background. This makes us think that there might be room for rethinking grade repetition practices in MS.

In these multilevel models, we separate the effects of socioeconomic background (ESCS) into within-school individual student effect and school composite effect. Both are negatively related to expected early leaving, but while the individual effect remains similar across samples and specifications, the between-school composite effect is much weaker after including other school-average composite indicators: school average achievement and school average likelihood of expected early leaving from education and training. The latter is consistently positively related to expected early leaving from education and training across different samples. Thus, while the compositional effect of student ESCS is associated with early leaving from education and training decisions, it is partly related to the compositional effects of the popularity of expected early leaving in a school. Relative to other schools with similar ESCS background, the likelihood of expected early leaving from education and training for an individual student increases if other students in the same school consider finishing school early. Obviously, this association is driven by school or student body characteristics that cannot be disentangled in large-scale surveys, but the strength of this relationship suggests that policies supporting schools with a large number of leavers are necessary to encourage individual students to continue education.

In contrast, the association with student achievement seems to be complex. At the individual level, higher achievement lowers the likelihood of early leaving. This is intuitive as students who do well are more likely to continue education. However, the compositional effect at the school level is positive, which means that in schools of similar socioeconomic background, with a similar share of students considering early leaving, and controlling for other characteristics in our model, higher average achievement appears to increase the likelihood of expected early

leaving for an individual student. One possible explanation is that in higher performing schools some students may be thinking about leaving education early as they perceive their own performance negatively, a phenomenon that could be similar to the big-fish-little-pond effect (Fang et al., 2018).

The multilevel regressions also include an individual level dummy for immigrant students and school-average share of immigrant students. Both are insignificant or negatively related to the possibility of early leaving after controlling for other compositional and individual effects, and also for school characteristics. Based on the available data, these results suggest that – after controlling for individual and compositional effects of other variables – we do not find evidence of a significant positive relationship between having a migrant background and declaring an intention to abandon formal education early, even in schools with more students having an immigrant background.

Finally, note that the models in columns (2), (4) and (6) explain almost all variance at the school level. Thus, the compositional effects specified in our regressions are almost entirely explaining variation in early leaving from education and training decisions across schools.

Table 8 provides additional results comparing associations between native and migrant students. Grade repetition and truancy have similar associations with native and immigrant students. In a fuller specification in columns (4) to (6), within- and between-school associations of socioeconomic background and expected early leaving from education and training are also similar across native and immigrant students. The peer effects of having high achieving students and colleagues who consider early leaving from education and training are also weaker for migrant students. The latter association is significantly lower in all three regressions for immigrant students. Overall, these additional results demonstrate small differences between native and migrant students, and when significant, compositional and individual effects that are positively associated with expected early leaving from education and training tend to be smaller for immigrant students.

Table 7.   **Multilevel models with additional student and composite predictors**

	PISA 2015		PISA 2018		2015 and 2018	
	(1)	(2)	(3)	(4)	(5)	(6)
Duration of early education and care	-0.00	0.00	0.13	0.07	0.03	0.01
	(0.04)	(0.04)	(0.08)	(0.07)	(0.04)	(0.03)
Sense of belonging	-0.00	0.01	0.09	0.14*	0.02	0.05
	(0.05)	(0.05)	(0.07)	(0.07)	(0.04)	(0.04)
Truancy	0.12**	0.08	0.26***	0.16***	0.18***	0.11***
	(0.04)	(0.04)	(0.05)	(0.05)	(0.03)	(0.03)
Grade repetition	1.04***	0.56*	0.36*	0.16	0.81***	0.42*
	(0.29)	(0.26)	(0.17)	(0.16)	(0.21)	(0.17)
Female	0.06	-0.00	-0.22	-0.20	-0.00	-0.03
	(0.11)	(0.11)	(0.14)	(0.13)	(0.09)	(0.10)
Within-school ESCS effect	-0.40***	-0.36***	-0.46***	-0.44***	-0.42***	-0.38***
	(0.05)	(0.05)	(0.08)	(0.09)	(0.04)	(0.05)
Between-school ESCS effect	-1.90***	-0.15*	-2.68***	-0.54*	-2.11***	-0.28*
	(0.15)	(0.07)	(0.39)	(0.26)	(0.18)	(0.11)
Migrant	-0.23	-0.49**	0.07	0.03	-0.14	-0.30*
	(0.14)	(0.15)	(0.21)	(0.22)	(0.12)	(0.13)
School share of migrant students	-0.08	-0.10	-2.59**	-1.07*	-1.06**	-0.40
	(0.29)	(0.18)	(0.84)	(0.45)	(0.37)	(0.21)
Student achievement (science or reading)		-0.01***		-0.00***		-0.01***
		(0.00)		(0.00)		(0.00)
School average achievement		0.01***		0.00*		0.01***
		(0.00)		(0.00)		(0.00)
School share of students expecting early leaving		5.95***		6.90***		6.16***
		(0.17)		(0.76)		(0.26)
2018 wave dummy					Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-4.35***	-3.02***	-4.40***	-3.94***	-4.12***	-3.25***
	(0.34)	(0.41)	(0.36)	(1.12)	(0.33)	(0.53)
Variance at the school level	0.74***	0.00*	1.74***	0.25*	1.12***	0.00
N of students	99 823	99 823	77 313	77 313	177 136	177 136

NB: Standard errors in parentheses; * p<0.05, ** p<0.01, *** p<0.001

Table 8. **Additional results with interaction terms for migrant students**

	2015	2018	2015 and 2018	2015	2018	2015 and 2018
	(1)	(2)	(3)	(4)	(5)	(6)
Truancy	0.12**	0.26***	0.18***	0.08	0.13*	0.09**
	(0.04)	(0.05)	(0.03)	(0.05)	(0.05)	(0.03)
Grade repetition	1.04***	0.35*	0.81***	0.45	0.24	0.37
	(0.30)	(0.17)	(0.22)	(0.30)	(0.19)	(0.20)
Female	0.06	-0.21	-0.00	-0.00	-0.20	-0.03
	(0.11)	(0.13)	(0.09)	(0.11)	(0.13)	(0.10)
Within-school ESCS effect	-0.44***	-0.48***	-0.46***	-0.39***	-0.46***	-0.41***
	(0.05)	(0.10)	(0.05)	(0.05)	(0.11)	(0.05)
Between-school ESCS effect	-1.86***	-2.66***	-2.08***	-0.15*	-0.52*	-0.28*
	(0.15)	(0.40)	(0.18)	(0.08)	(0.26)	(0.12)
Migrant	-0.32*	0.04	-0.20	-0.79***	0.07	-0.44*
	(0.15)	(0.25)	(0.13)	(0.22)	(0.25)	(0.17)
School share of migrant students	-0.21	-2.64**	-1.15**	-0.41*	-0.96*	-0.55*
	(0.31)	(0.88)	(0.38)	(0.20)	(0.49)	(0.22)
Migrant*ESCS	0.23	0.09	0.18	0.16	0.10	0.12
	(0.13)	(0.16)	(0.10)	(0.14)	(0.18)	(0.11)
Migrant*school average ESCS	-0.54**	-0.15	-0.39*	-0.20	-0.12	-0.16
	(0.20)	(0.36)	(0.18)	(0.25)	(0.35)	(0.21)
Migrant*truancy				0.03	0.16	0.09
				(0.11)	(0.13)	(0.08)
Migrant*repeat				0.66	-0.34	0.26
				(0.37)	(0.41)	(0.27)
Within-school achievement				-0.01***	-0.00***	-0.01***
				(0.00)	(0.00)	(0.00)
School average achievement				0.01***	0.00*	0.01***
				(0.00)	(0.00)	(0.00)
Migrant*school average achievement				-0.00	-0.01*	-0.00*
				(0.00)	(0.00)	(0.00)
School average share of expected early leavers from education and training				5.98***	6.93***	6.18***
				(0.17)	(0.78)	(0.26)
Migrant*school share of leavers				-1.56***	-2.15*	-1.72***
				(0.46)	(1.00)	(0.45)
2018 wave dummy			Yes			Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-4.39***	-3.94***	-4.05***	-3.07***	-3.83***	-3.26***
	(0.37)	(0.25)	(0.33)	(0.39)	(1.08)	(0.51)
Variance at the school level	0.73***	1.74***	1.12***	0.00**	0.26*	0.00
N of students	99823	77313	177136	99823	77313	177136

NB: Standard errors in parentheses; * $p<0.05$, ** $p<0.01$, *** $p<0.001$

CHAPTER 5.
Conclusions

This paper has considered expectations towards early leaving from education and training among 15-year-olds in the European Union, in particular distinguishing between native and immigrant students. Education and skills differences are seen in the literature as fundamental factors for economic divides (Hanushek and Wößmann, 2015). Early school leaving is one of the reasons behind unemployment and a generally weaker position on the labour market for individuals, and also for disparities in social and political life (Cedefop 2016a; 2016b; 2016c; 2016d). A better understanding of education divergences from a European perspective, and subsequently improving educational attainment, may help in countering widening divides within and across MS.

The recent refugee crisis has brought the issue of immigrants more in focus in the EU agenda (European Commission, 2016), while migration patterns among EU members have led to debates in both origin and destination countries. The continuing war in Ukraine has also now resulted in a new scale of refugee flows, with millions of Ukrainians fleeing their country within a few days and weeks of the beginning of the war.

At the same time, the EU target of reducing the number of early school leavers to less than 9% by 2030 shows the importance that policy-makers have given to the reduction in ELET. ELET reflects and reinforces a range of individual and social disadvantages, which have negative consequences for students, society and the State.

In this study, we have used OECD's most recent PISA 2015 and 2018 data to explore the importance of immigrant status – together with other factors measured at the school and individual level – in accounting for the likelihood of expected early leaving from education and training. The data most used for measuring educational achievement is provided by PISA, whose data on migrants are available for the majority of MS. However, in six EU MS, the sample size of immigrant students is too low for reliable comparisons, or information about schooling expectations is not available. Our results are based on around 200 000 15-year-old students from 21 EU MS, including all MS with a considerable population of students with an immigrant background.

We analyse expected early leaving from education and training in PISA in various ways. First, we present descriptive statistics to get a better intuition for the data and for checking for consistency between the 2015 and 2018 surveys. We

then run a set of two-level multilevel logit models including both the student and the school level, controlling for the country fixed effects. The explanatory variables employed follow the standard OECD methodology but we also constructed a new variable on truancy from PISA responses. The final models compare associations of selected variables for immigrant and native students and decompose the effects into student and school level. In our full specification most of the variation at the school level is explained by individual and school-composite predictors.

Overall, we find that in a variety of settings, once we control for individual and school characteristics, immigrant status does not have a significant association with the expected likelihood of early leaving from education and training. The latter is rather associated with other factors, that are unequally distributed between migrants and natives, such as socioeconomic status and achievement. Moreover, for immigrant students, socioeconomic status is found to have a similar or weaker relationship with the likelihood of expected early leaving. Further, according to our results, the average socioeconomic status in a school appears to have a strong relationship with the likelihood of expected early leaving, meaning that students in schools with many disadvantaged peers are more likely to consider early leaving from education and training. However, for students with an immigrant background this association seems to be weaker, or at least similar, as in some specifications the estimated coefficient is insignificant. Finally, while the share of students considering early leaving from education and training is one of the strongest predictors of an individual student considering early leaving, this association is also found to be weaker for students with an immigrant background, even after controlling for similar differences in the associations of socioeconomic background, achievement and other variables.

Students with an immigrant background seem to be similar to native students (i.e. there is no significant association with immigration status) or even more resilient, at least when considering their expectations, despite difficult circumstances and negative peer associations in schools with disadvantaged students. Other factors that we considered in this research have a similar association with early leaving from education and training considerations for native and immigrant students, but are worth discussing as the relationships are robust to model specifications and different samples. Grade repetition and truancy, i.e. skipping classes or being regularly late to school, appear to be both positively related to the likelihood of expected early leaving from education and training. All students are found to be similarly negatively affected by these two factors. Grade repetition is still a popular practice in some EU MS and we provide additional evidence on its detrimental education consequences. Truancy should be seen as a sign of behaviour that could lead to early leaving from education and training and

approached by practices and policies that encourage learning rather than punish students who are already close to finishing education early. While males have usually higher ELET rates, in our study we do not find that gender is a significant factor for expectations of early leaving across Europe when a whole set of explanatory variables is included. Also, a sense of belonging, i.e. the attachment students feel to their peers and school, appears not to be related to expected early leaving from education and training.

This study overcomes the main limitation of national research based on representative surveys. It provides results based on a large sample size as it pools data from all EU MS with a substantial number of students with an immigrant background. The results suggest that, on average, migrant students are similar to native students or are even more eager to continue their education, after taking into account individual and school compositional effects. Thus, early leaving from education and training policies should focus not on migrant students but more generally on disadvantaged students and schools with a large share of early leavers. Discipline and avoiding ineffective policies that discourage learning, like grade repetition, seem to be crucial for limiting early leaving from education and training in EU MS.

References

[URLs accessed 21.9.2022]

Benhenda, A. and Grenet, J. (2015*). How much does grade repetition in French primary and secondary schools cost?* IPP Note, No 17. Paris: Institut des Politiques Publiques. https://www.ipp.eu/en/publication/n17-how-much-does-grade-repetition-in-french-primary-and-secondary-schools-cost/

Blanden, J. and McNally, S. (2015). *Reducing inequality in education and skills: implications for economic growth*. EENEE analytical report, No 21. https://eenee.eu/en/resources/library/reducing-inequality-in-education-and-skills-implications-for-economic-growth/

Brinbaum, Y. and Cebolla-Boado, H. (2007). The school careers of ethnic minority youth in France: success or disillusion? *Ethnicities*, Vol. 7, No 3, pp. 445-474.

Brunello, G. and De Paola, M. (2013). *The costs of early school leaving in Europe*. EENEE analytical report, No 17 prepared for the European Commission. https://eenee.eu/en/resources/library/the-costs-of-early-school-leaving-in-europe/

Cedefop (2016a). *Leaving education early: putting vocational education and training centre stage, volume II: evaluating policy impact, executive summary*. https://www.cedefop.europa.eu/files/5558_en_executive_summary.pdf

Cedefop (2016b). *Leaving education early: putting vocational education and training centre stage, volume I: investigating causes and extent, executive summary.* https://www.cedefop.europa.eu/files/5557_en_executive_summary.pdf

Cedefop (2016c). *Leaving education early: putting vocational education and training centre stage, volume II: evaluating policy impact*. Luxembourg: Publications Office. Cedefop research paper, No 58. http://dx.doi.org/10.2801/967263

Cedefop (2016d). *Leaving education early: putting vocational education and training centre stage, volume I: investigating causes and extent*. Luxembourg: Publications Office. Cedefop research paper, No 57. http://dx.doi.org/10.2801/893397

Cedefop (2021). Labour market integration of refugees: practices on recognition of skills and qualifications. https://www.cedefop.europa.eu/en/events/labour-market-integration-refugees-practices-recognition-skills-qualifications

Cedefop (2022a). *VET toolkit for tackling early leaving*. https://www.cedefop.europa.eu/en/tools/vet-toolkit-tackling-early-leaving/about-toolkit

Cedefop (2022b). *VET toolkit for empowering NEETs.* https://www.cedefop.europa.eu/en/tools/neets/about-toolkit

Cedefop (2022c). *VET toolkit for tackling early leaving: identify.*
https://www.cedefop.europa.eu/en/tools/vet-toolkit-tackling-early-leaving/identify

Council of the European Union (2021). Council Resolution on a strategic framework for European cooperation in education and training towards the European Education Area and beyond (2021-30). *Official Journal of the European Union*, C 66, 26.2.2021, pp. 1-21. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32021G0226(01)

De Paola, M. and Brunello, G. (2016). *Education as a tool for the economic integration of migrants.* EENEE analytical report, No 27. https://data.europa.eu/doi/10.2766/46295

Dustmann, C.; Frattini, T. and Lanzara, G. (2012). Educational achievement of second-generation immigrants: an international comparison. *Economic Policy*, Vol. 27, No 69, pp. 143-185.

Dustmann, C. and Glitz, A. (2011). Migration and education. In: *Handbook of the Economics of Education*, Vol. 4, chapter 4, pp. 327-439.

Dustmann, C. and Theodoropoulos, N. (2010). Ethnic minority immigrants and their children in Britain. *Oxford Economic Papers*, Vol. 62, No 2, pp. 209-233.

European Commission (2013). *Reducing early school leaving: key messages and policy support: final report of the Thematic Working Group on Early School Leaving.* https://education.ec.europa.eu/sites/default/files/early-school-leaving-group2013-report_en.pdf

European Commission (2016). *European Semester thematic factsheet: early school leavers.* https://ec.europa.eu/info/sites/info/files/european-semester_thematic-factsheet_early-school-leavers_en.pdf

European Commission (2017). *Strengthening European identity through education and culture: the European Commission's contribution to the Leaders' meeting in Gothenburg, 17 November 2017. COM(2017) 673 final.* https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2017%3A673%3AFIN

European Commission. (2021). *Education and training monitor 2021: education and well-being.* Luxembourg: Publications Office. https://data.europa.eu/doi/10.2766/743550

European Commission et al. (2014). *Tackling early leaving from education and training in Europe: strategies, policies and measures.* Luxembourg: Publications Office. https://data.europa.eu/doi/10.2797/33979

Eurostat. (2021). *Early leavers from education and training by sex and labour status* [edat_lfse_14]. https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=edat_lfse_14&lang=en

Fang, J. et al. (2018). The big-fish-little-pond effect on academic self-concept: a meta-analysis. *Frontiers in psychology*, 9, 1569. https://doi.org/10.3389/fpsyg.2018.01569

Feliciano, C. (2006). Beyond the family: the influence of premigration group status on the educational expectations of immigrants' children. *Sociology of Education*, Vol. 79, No 4, pp. 281-303.

Flisi, S.; Meroni, E.C. and Vera-Toscano, E. (2016). *Educational outcomes and immigrant background*. Luxembourg: Publications Office. JRC technical report. https://data.europa.eu/doi/10.2791/026577

Gabay-Egozi, L.; Shavit, Y. and Yaish, M. (2014). Gender differences in fields of study: the role of significant others and rational choice motivations. *European Sociological Review*, Vol. 31, No 3, pp. 284-297.

Goldenberg, C. et al. (2001). Cause or effect? A longitudinal study of immigrant Latino parents' aspirations and expectations, and their children's school performance. *American Educational Research Journal*, Vol. 38, No 3, pp. 547-582.

Gutierrez, O.D.M. and Lopez-Agudo, L.A. (2016). On the endogeneity of parental expectations and children's academic attainment. *Investigaciones de Economía de La Educación*, Vol. 11, No 11, pp. 551-568.

Hagelskamp, C.; Suárez-Orozco, C. and Hughes, D. (2010). Migrating to opportunities: how family migration motivations shape academic trajectories among newcomer immigrant youth. *Journal of Social Issues*, Vol. 66, No 4, pp. 717-739.

Hanushek, E.A. and Wößmann, L. (2006). Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *The Economic Journal*, Vol. 116, No 510, pp. 63-76. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0297.2006.01076.x

Hanushek, E.A. and Wößmann, L. (2015*). The knowledge capital of nations: education and the economics of growth.* MIT Press.

Heath, A. and Brinbaum, Y. (2007). Guest editorial: explaining ethnic inequalities in educational attainment. *Ethnicities*, Vol. 7, No 3, pp. 291-305. https://journals.sagepub.com/doi/10.1177/1468796807080230

Heath, A.; Rothon, C. and Kilpi, E. (2008). The second generation in Western Europe: education, unemployment, and occupational attainment. *Annual Review of Sociology,* Vol. 34, pp. 211-235. https://www.annualreviews.org/doi/10.1146/annurev.soc.34.040507.134728

Rabe-Hesketh, S. and Skrondal, A. (2006). Multilevel modelling of complex survey data. *Journal of the Royal Statistical Society Series A*, Vol. 169, No 4, pp. 805-827.

Hippe, R.; Araújo, L. and Dinis da Costa, P. (2016). *Equity in education in Europe.* Luxembourg: Publications Office. JRC technical report. https://data.europa.eu/doi/10.2791/255948

Hippe, R.; Jakubowski, M. and Araújo, L. (2018). *Regional inequalities in PISA: the case of Italy and Spain.* Luxembourg: Publications Office. JRX technical report. https://data.europa.eu/doi/10.2760/495702

Ikeda, M. and García, E. (2014). Grade repetition: a comparative study of academic and non-academic consequences. *OECD Journal: Economic Studies*, Vol. 1, pp. 269-315. https://www.oecd-ilibrary.org/economics/grade-repetition_eco_studies-2013-5k3w65mx3hnx

Jakubowski, M. and Pokropek, A. (2015). Reading achievement progress across countries. *International Journal of Educational Development*, Vol. 45, pp. 77-88.

Jonsson, J.O. and Rudolphi, F. (2010). Weak performance - strong determination: school achievement and educational choice among children of immigrants in Sweden. *European Sociological Review*, Vol. 27, No 4, pp. 487-508.

Kao, G. and Tienda, M. (1995). Optimism and achievement: the educational performance of immigrant youth. *Social Science Quarterly*, Vol. 76, No 1, pp 1-19.

Kristen, C. and Dollmann, J. (2010). Sekundäre Effekte der ethnischen Herkunft: Kinder aus türkischen Familien am ersten Bildungsübergang. In: Becker, B. and Reimer, D. *Vom Kindergarten bis zur Hochschule: die Generierung von ethnischen und sozialen Disparitäten in der Bildungsbiographie,* pp. 117–144.

Lüdemann, E. and Schwerdt, G. (2010). Migration background and educational tracking: is there a double disadvantage for second-generation immigrants? *Journal of Population Economics*, Vol. 26, pp. 455-481.

Massey, D.S. and Taylor, J.E. (2004). *International migration: prospects and policies in a global market*. Oxford: Oxford University Press.

Minello, A. (2014). The educational expectations of Italian children: the role of social interactions with the children of immigrants. *International Studies in Sociology of Education*, Vol. 24, No 2, pp. 127-147.

Minello, A. and Barban, N. (2012). The educational expectations of children of immigrants in Italy*. The Annals of the American Academy of Political and Social Science*, Vol. 643, No 1, pp. 78-103.

Murat, M. (2011). *Do immigrant students succeed?: evidence from Italy and France based on PISA 2006*. Università degli Studi di Modena e Reggio Emilia, Dipartimento di Economia Politica.
http://core.ac.uk/download/pdf/6220431.pdf

Nauck, B. and Kohlmann, A. (1999). Kinship as social capital: network relationships in Turkish migrant families. In: Supper, S and Richter, R. (eds.). *New Qualities in the lifecourse. intercultural aspects*, pp. 199-218.

OECD (2015a). *Helping immigrant students to succeed at school and beyond.* https://www.oecd.org/education/Helping-immigrant-students-to-succeed-at-school-and-beyond.pdf

OECD (2015b). *Immigrant students at school: easing the journey towards Integration*. Paris: OECD Publishing.
http://dx.doi.org/10.1787/9789264249509-en

OECD (2016a). PISA 2015 results (volume I): excellence and equity in education. Paris: OECD Publishing http://dx.doi.org/10.1787/9789264266490-en

OECD (2017a). PISA 2015 results (volume III): students' well-being. Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264273856-en

OECD (2017b). PISA 2015: technical report: chapter 16: procedures and construct validation of context questionnaire data. http://www.oecd.org/pisa/data/2015-technical-report/

OECD and European Union. (2015). Indicators of immigrant integration 2015: settling in. http://dx.doi.org/10.1787/9789264234024-en

Phalet, K.; Andriessen, I. and Lens, W. (2004). How future goals enhance motivation and learning in multicultural classrooms. *Educational Psychology Review,* Vol. 16, No 1, pp. 59-89.

Portes, A. et al. (2010). Moving ahead in Madrid: aspirations and expectations in the Spanish second generation. *International Migration Review*, Vol. 44, No 4, pp. 767-801.

Psifidou, I. and Treves, A. (2022). *Making VET inclusive for Ukrainian refugee students.* https://www.cedefop.europa.eu/en/blog-articles/making-vet-inclusive-ukrainian-refugee-students

Relikowski, I.; Yilmaz, E. and Blossfeld, H.-P. (2012). Wie lassen sich die hohen Bildungsaspirationen von Migranten erklären? Eine Mixed-Methods-Studie zur Rolle von strukturellen Aufstiegschancen und individueller Bildungserfahrung. In: Becker, R. and Solga, H. (eds). *Soziologische Bildungsforschung,* pp. 111-136.

Rodrigues, M. (2018). *Can digital technologies help reduce the immigrant-native educational achievement gap?* Luxembourg: Publications Office. JRC technical report. https://data.europa.eu/doi/10.2760/953261

Ruhose, J. and Schwerdt, G. (2016). Does early educational tracking increase migrant-native achievement gaps? Differences-in-differences evidence across countries. *Economics of Education Review*, Vol. 52, pp. 134-154. https://www.sciencedirect.com/science/article/abs/pii/S0272775715300741

Salikutluk, Z. (2016). Why do immigrant students aim high? Explaining the aspiration-achievement paradox of immigrants in Germany. *European Sociological Review*, Vol. 32, No 5, pp. 581-592.

Schneeweis, N. (2015). Immigrant concentration in schools: consequences for native and migrant students. *Labour Economics*, Vol. 35, pp. 63-76.

Schnepf, S. V. (2007). Immigrants' educational disadvantage: an examination across ten countries and three surveys. *Journal of Population Economics*, Vol. 20, No 3, pp. 527-545.

Sewell, W.H.; Haller, A.O. and Ohlendorf, G.W. (1970). The educational and early occupational status attainment process: Replication and revision. *American Sociological Review*, Vol. 36, No 6, pp. 1014-1027.

Sewell, W. H. and Hauser, R. M. (1972). Causes and consequences of higher education: Models of the status attainment process. *American Journal of Agricultural Economics*, Vol. 54, No 5, pp. 851–861.

Sewell, W. H. and Shah, V. P. (1968). Social class, parental encouragement, and educational aspirations*. American Journal of Sociology*, Vol. 73, No 5, pp. 559–572.

Sikora, J. and Biddle, N. (2015). How gendered is ambition? Educational and occupational plans of Indigenous youth in Australia. *International Journal of Educational Development*, Vol. 42, pp. 1–13.

Teney, C.; Devleeshouwer, P. and Hanquinet, L. (2013). Educational aspirations among ethnic minority youth in Brussels: does the perception of ethnic discrimination in the labour market matter? A mixed-method approach. *Ethnicities*, Vol. 13, No 5, pp. 584-606.

Tjaden, J. D. and Hunkler, C. (2017). The optimism trap: migrants' educational choices in stratified education syste

Annex 1.
Definitions of immigrant background

Non-immigrant (native): born in the country of assessment/survey, or with at least one parent born in that country; born abroad with at least one parent born in the country of assessment/survey.

Second-generation immigrants: born in the country of assessment/survey, but whose parents were born in another country. In this case, it is not the individuals who move, but their parents (definition based on previous generation's move).

First-generation immigrants: born outside the country of assessment/survey and whose parents were also born in another country.

Annex 2.
Details on the questions asked in PISA to construct the variables

Below, we provide additional information on the variables included in this study, taken directly from, and as presented by, OECD sources (2017a, 2017b), except for truancy, which we constructed ourselves (see main text for more details).

Expected leaving from education and training

ST111Q01TA	Which of the following do you expect to complete?
1	<ISCED level 2>
2	<ISCED level 3B or C>
3	<ISCED level 3A>
4	<ISCED level 4>
5	<ISCED level 5B>
6	<ISCED level 5A or 6>

Terms enclosed in brackets < > in the following descriptions were replaced in the national versions of the student, school and parent questionnaires by the appropriate national equivalent. For example, the term <qualification at ISCED level 5A> was translated in the United States into 'bachelor's degree, postgraduate certificate program, master's degree program or first professional degree program'.

Immigration background

The PISA database contains three country-specific variables relating to the students' country of birth, their mother and father (COBN_S, COBN_M, and COBN_F). The items ST019Q01TA, ST019Q01TB and ST019Q01TC were recoded into the following categories: (1) country of birth is the same as country of assessment and (2) other. The index of immigrant background (IMMIG) was calculated from these variables with the following categories: native students (those students who had at least one parent born in the country), (2) second-generation students (those born in the country of assessment but whose parent(s) were born in another country) and (3) first-generation students (those students born outside the country of assessment and whose parents were also born in another country). Students with missing responses for either the student or for both parents were assigned missing values for this variable.

Item	Country of birth
COBN_F	Country of Birth National Categories- Father
COBN_M	Country of Birth National Categories- Mother
COBN_S	Country of Birth National Categories- Self

IMMIG	Index Immigration status
1	Native
2	Second-Generation
3	First-Generation

Truancy

An index which we constructed using the responses to the test questions below:

Item	In the last 2 full weeks of school, how often did the following things occur?
ST062Q01TA	In the last 2 full weeks of school, how often: I <skipped> a whole school day
ST062Q02TA	In the last 2 full weeks of school, how often: I <skipped> some classes
ST062Q03TA	In the last 2 full weeks of school, how often: I arrived late for school

Early childhood education and care

Questions ST125 and ST126 measure the starting age in ISCED 1 and ISCED 0. The indicator DURECEC is built as the difference of ST126 and ST125 plus the value of '2' to indicate the number of years a student spent in early childhood education and care.

Grade repetition

The grade repetition variable (REPEAT) was computed by recoding variables ST127Q01TA, ST127Q02TA, and ST127Q03TA. REPEAT took the value of '1' if the student had repeated a grade in at least one ISCED level and the value of '0' if 'no, never' was chosen at least once, given that none of the repeated grade categories were chosen. The index is assigned a missing value if none of the three categories were ticked in any levels.

Item	Grade repetition
ST127Q01TA	Have you ever repeated a <grade>? At <ISCED 1>
ST127Q02TA	Have you ever repeated a <grade>? At <ISCED 2>
ST127Q03TA	Have you ever repeated a <grade>? At <ISCED 3>

Programme orientation

Programme orientation (ISCEDO) indicates whether the programme's curricular content was general, pre-vocational or vocational. Note: 'modular' also exists as a programme title in Luxembourg.

ISCED level

Programme level (ISCEDL) indicates whether students were at the lower or upper secondary level (ISCED 2 or ISCED 3).

Programme designation

Programme designation (ISCEDD) indicates the designation of the study programme: (1) 'A' (general programmes designed to give access to the next programme level); (2) 'B' (programmes designed to give access to vocational studies at the next programme level); (3) 'C' (programmes designed to give direct access to the labour market); or (4) 'M' (modular programmes that combine any or all of these characteristics).

Sense of belonging

PISA 2015 asked students about their sense of belonging to school (ST034) using six trend items previously used in PISA 2012 (ID in 2012: ST87). The answering format was a four-point Likert scale with the answering categories 'strongly agree', 'agree', 'disagree', and 'strongly disagree'; the derived IRT-Scale is named BELONG. Items ST034Q02TA, ST034Q03TA and ST034Q05TA were reverse-coded so that higher WLEs and higher difficulty correspond to a higher level of sense of belonging on all items.

Age at arrival

ST021Q01TA	How old were you when you arrived in <country of test>?
1	age 0 - 1
2	age 1
3	age 2
4	age 3
5	age 4
6	age 5
7	age 6
8	age 7
9	age 8
10	age 9
11	age 10

ST021Q01TA	How old were you when you arrived in <country of test>?
12	age 11
13	age 12
14	age 13
15	age 14
16	age 15
17	age 16
95 / .V	Valid Skip
97 / .N	Not applicable
98 / .I	Invalid
99 / .M	No response
System missing	Missing

Language at home

Students indicated what language they usually speak at home (ST022), and the database includes a derived variable (LANGN) containing a country-specific code for each language. In addition, an internationally comparable variable was derived from this information with the following categories: (1) language at home is the same as the language of assessment for that student and (2) language at home is another language.

At the student level we include the following variables:

Truancy measures whether students are skipping school days, skipping classes or arriving late at school. In fact, students were asked several questions about truancy. We used their responses to estimate an index of truancy using the principal component analysis. The intuition is that skipping classes may have a negative association with performance and thus a positive one with the expectation of leaving early.

Economic, social and cultural status (ESCS) is an OECD index measuring student socioeconomic background. PISA measures ESCS with an extensive set of questions related to parent occupation, education and household cultural, education and economic resources. It is usually positively associated with PISA scores. The index is standardised to have a mean of 0 and a standard deviation of 1 across OECD countries (weighting each country equally). A lower socioeconomic background of students generally increases early leaving rates, as has been shown by a multitude of studies on this subject (European Commission et al., 2014).

Immigrant background refers to first- or second-generation immigrant students. As first and second-generation students have different immigration

backgrounds, the results may, in many cases, vary according to each group. In consequence, we include this distinction among immigrant students.

Language spoken at home controls for the fact that the language taught at school and the language spoken at home may be different. Thus, a student may have more difficulty with subjects at school and in socialising with schoolmates when the same language is not practised in both home and school environments. Thus, not speaking the language at home has also been shown to be negatively associated with PISA scores. In MS such as the Czech Republic, Slovenia, Finland and Sweden, the share of first-generation students who do not speak the same language at home as in school is more than 80%, while in Croatia this is true only for less than 10% (OECD, 2015a).

Sense of belonging shows how psychologically well students feel integrated and feel they belong to the school they attend. This is an important measure of the social integration of immigrant students. The OECD concludes that the sense of belonging varies widely among EU MS. For example, in the UK newly arriving (first-generation) immigrant students have a higher feeling of belonging than natives and second-generation students. Both of the latter two groups have a similar level of belonging. However, in France for example, second-generation students have the lowest sense of belonging of all groups. In contrast, integration is more progressive in Spain, Italy and Sweden. In these MS, second-generation students have similar (high) values to natives, while first-generation students have lower values (OECD, 2015a).

Age at arrival provides information on when an immigrant student has arrived in the destination country in which she took the PISA test. The intuition for including this variable is that it is easier for a student to learn and integrate into a new cultural and linguistic environment the younger she is (OECD, 2016b). The OECD also calls this the late-arrival penalty: '[i]n most OECD countries, immigrant students who arrived at the age of 12 or older – and have spent at most 4 years in their new country – lag farther behind students in the same grade in reading proficiency than immigrants who arrived at younger ages' (OECD, 2015a, p. 10).

Grade repetition is a policy that is common in some European countries and has been shown to negatively influence leaving from education and training (European Commission et al., 2014). Previous research using PISA data suggests that repeating a grade is not associated with improved student performance and shows a negative association with student attitudes (Ikeda and García, 2014). At the same time, this policy is very costly (Benhenda and Grenet, 2015). In PISA 2015, and in line with findings from previous PISA rounds, students who have repeated a grade at least once have lower scores (OECD, 2016b).

Finally, *gender* is a dummy variable, taking one for female students. Research has shown that girls may have different educational expectations than boys (e.g. Sikora and Biddle, 2015). For example, girls' educational expectations are much higher in Italy (Minello, 2014). In general, girls also have lower realised early leaving rates than boys at the European level (European Commission, 2016; European Commission et al., 2014).

In some specifications, we also include students' *science scores*.[15] In fact, science achievement is the main domain in PISA 2015. This means that science was measured with the highest possible precision and every student taking the 2015 PISA test answered a number of science-related test items. While reading and mathematics were also measured in 2015, the number of test items for these domains was smaller and the resulting measurement is less precise. Thus, we decided to include science performance and not reading or mathematics performance. In PISA, science is defined as 'the ability to engage with science-related issues, and with the ideas of science, as a reflective citizen. A scientifically literate person is willing to engage in reasoned discourse about science and technology, which requires the competencies to explain phenomena scientifically, evaluate and design scientific enquiry, and interpret data and evidence scientifically' (OECD, 2016, p. 13). We use this variable to proxy for unobserved ability.

In addition, we consider the following variables at the school level:

Concentration of immigrant students in schools indicates the relative prevalence of immigrant students in schools (averaged for schools). While there can be some advantages of arriving in a school with students from a similar cultural background, in many cases the concentration of immigrants is related to a concentration of socioeconomically disadvantaged students (OECD, 2015a). In PISA 2015, a high concentration of immigrant students in schools is not associated with poorer student performance, while, individually, immigrant students on average show lower performance in EU MS (OECD, 2016b).

We also include programme designation, which provides further information on the study programme (general or vocational level giving access to the next level, giving direct access to the labour market and to modular programmes). Similarly, programme orientation indicates whether students are enrolled in general, pre-

---

([15]) Student achievement is reflected by a set of so-called plausible values that reflect student outcomes and allow for estimating measurement error. In PISA 2015 and 2018, 10 plausible values are provided in the data sets and the analysis should replicate every estimation 10 times with each plausible value. The results averaged across 10 replications provide unbiased estimates of student achievement. In the calculation of standard errors, variation across 10 replications is included using special formulas that add estimates of measurement error to the estimates of sampling error.

vocational or vocational programmes. It should be noted that in many MS included in this analysis all 15-year-olds are in academic or general schools. Thus, the results should be interpreted cautiously. However, this variable provides interesting insights for some MS with a sufficient number of students in vocational education. Early tracking, before the age of 15, has been shown to increase education inequality (Ruhose and Schwerdt, 2016). However, at the country level, MS 'with a relatively weak VET system tend to have a higher problem of early leaving, likely due to the lack of sufficiently attractive non-academic programmes' (European Commission et al., 2014, p. 13). This has been confirmed by Cedefop, stating 'early leaving […] seems to be more acute in VET than in general education' (Cedefop, 2016a, p. 6). Still, it also needs to be taken into account that 'VET's more practical approach, particularly when work-based, can make learning more meaningful for some young people and play a role in motivating them to continue in, or return to, education and training' (Cedefop, 2016a, p. 6). In addition, VET can be a safety net for those dropping out from general education and who could have otherwise been early leavers (Cedefop, 2016b, p. 4).

We also control for the *ISCED level*. In fact, PISA uses the age of students as criteria for inclusion in the survey, i.e. only 15-year-olds are allowed to participate. As a consequence, some students are already in an ISCED 3 year, and many others are still in ISCED 2. While this does not mean that they may expect to complete ISCED 3, it may potentially influence the results. Therefore, we control for this factor in the regression analysis ([16]).

Finally, we also include school means for the ESCS, expected early leaving from education and training and science scores. In other words, the student's data were aggregated and averaged for schools to be able to control for further school environment effects.

([16]) See also details on the minimum age during compulsory schooling in the appendix.

EXPECTED EARLY LEAVING AMONG NATIVE AND MIGRANT STUDENTS: EVIDENCE FROM PISA FOR EU MEMBER STATES

Early leavers from education and training become generally disadvantaged socially and economically in later stages in life, so it is important to understand better the motivations for early leaving and provide adequate policy solutions. This study analyses the factors that are most strongly related to the likelihood of leaving education and training early. It places special attention on migrant status, given the recent migration waves within and outside Europe, which are further exacerbated by the continuing war in Ukraine. To this end, we use the OECD most recent PISA data for 2015 and 2018 in a number of two-level logit regression models, including student- and school-level variables. The results show that migrant students do not differ structurally from EU natives in their likelihood of expected early leaving, implying that it is more important to focus on common factors.