CPB Netherlands Bureau for
Economic Policy Analysis

Can skills differences explain the gap in track recommendation by socio-economic status?

In the Netherlands students are tracked relatively early, at the end of primary school. Students from lower socio-economic backgrounds receive significantly lower track recommendations. Around 55% of the difference in advice can be explained by differences in cognitive and non-cognitive skills. Additionally controlling for regional, school and family background still leaves around 30% of the difference unexplained.

A reevaluation of the recommendation for students who outperform their initial recommendation in a nationwide test does not lead to a closing of the initial gap.

CPB Discussion Paper

Maria Zumbuehl, Nihal Chehber, Rik Dillingh

May 2022

Doi: https://doi.org/ 10.34932/sv55-2d58

Can skills differences explain the gap in track recommendation by socio-economic status?

Maria Zumbuehl, Nihal Chehber, Rik Dillingh*

May 18, 2022

JEL codes: I21, I24, J24
Economics of education; tracking; cognitive skills; non-cognitive skills

Abstract

Tracking early in the school career can influence a student's further educational path significantly. We study the track advice at the end of primary school in the Netherlands, where teachers give a track advice based on a student's previous performance and their impression of the student's ability. If the student outperforms the initial advice in the subsequent nationwide test, the school reevaluates the student and can — but does not have to — update the final advice. We use cognitive and non-cognitive skills measurements that are collected three years before the tracking decision is made, linked with the teachers initial and revised advice, as well as background information from register data. We find that with equal skills, students from lower socio-economic backgrounds receive on average lower advice, while students with a migration background receive on average higher advice. A decomposition of the total difference in initial advice between students from high versus low educated parents shows that around 55% of the difference in advice can be explained by differences in cognitive and non-cognitive skills. Adding additional information about the family, school and place of residence, we can explain about 71% of the difference between students with low and high educated parents. We do not find a significant change in the gap in advice between children from different socio-economic backgrounds after the nationwide test and reevaluation procedure.

*We thank Lex Borghans, Kimberley Lek, Dinand Webbink and the participants of the CPB research seminar for helpful comments and suggestions. This project has been fundend by the Dutch Ministry of Education (OCW), and Ministry of Social Affairs (SZW).

1 Introduction

Tracking early in the school career has an impact on students' skills development, their later educational achievements and labor market outcomes.[1] In particular lower performing children have persistently been found to be negatively affected by early tracking, while there is mixed evidence on the positive effects for children on the upper end of the skills distribution.[2] Since children with a lower socio-economic background are more frequently in the group of students who face a disadvantage through this system, the tracking also increases the difference in outcomes between groups.[3]

The mechanism through which the placement of students is determined may add another layer to this increase in inequality of opportunities. Ideally the placement into tracks should be based solely on a student's ability. However, since pure ability is not observable, frequently the placement is based on achievement tests, grade point average and/or an assessment of the students' abilities by their teachers. All of these potential inputs can be related to the socio-economic background of children and hence widen the gap if taken into consideration for the students placement. In particular teacher assessment might be biased, if teachers judge skills and achievements differently for specific groups of students[4] or base their decisions on additional factors, that correlate with the socio-economic background of students. Falk et al. (2020), for example, find a strong relation between socio-economic status and the placement of students conditional on the students' skills. They find that an intervention through a low-intensity mentoring program can mitigate the relationship between disadvantaged background and low track placement.

In this study we investigate the relation between track recommendation and socio-economic background in the Netherlands. In a recent report on behalf of the Dutch Ministry of Education, Geven et al. (2018) show that students from higher socio-economic groups receive a better evaluation from teachers than their peers from lower

[1] See e.g. Korthals et al. (2021) for the development of cognitive and non-cognitive skills, Borghans et al. (2019) or Van Elk et al. (2011) for education outcomes and Canaan (2020) for labor market outcomes.

[2] Borghans et al. (2020) and Roller and Steinberg (2020) find negative effects on lower skilled students and (partially) positive effects for higher skilled children in systems where all children are tracked. In a system where only high-ability students are tracked, Card and Giuliano (2016) find positive effects for the tracked students and no averse effects for the not-tracked students.

[3] Pekkala Kerr et al. (2013) report a decrease in test differences by parental education after the tracking had been delayed by two years, while Sulzmaier (2020) reports an increase in intergenerational transmission of educational outcomes after a decrease in the tracking age.

[4] See e.g. Carlana (2019), Sprietsma (2013) or Hanna and Linden (2012) for differences in grading, and Timmermans et al. (2015) or Papageorge et al. (2020) for differences in expectations.

socio-economic groups with similar academic performance. We study this gap, by conditioning not only on previous performance but also non-school-related capacity test scores and non-cognitive skills. We first investigate the determinants of the initial school advice, to learn more about the factors that contribute to the teacher's assessment. We then study how much of the difference in advice by socio-economic status we can explain by differences in skills and other observable characteristics.

The Dutch system tracks students relatively early at age 12 and is characterized by a strongly stratified tracking system with nine possible track recommendation categories. This allows for a more nuanced investigation of the process that leads to the teacher's track recommendation. The track advice is mainly based on the teacher's assessment of the student, however the precise process and inputs to this assessment are not regulated by the education system. After the initial advice is given, students participate in a nationwide test. If they outperform their initial advice, the school has to reevaluate the advice, but it is up to the school if this reevaluation leads to a readjustment of the initial advice or not. We use this two stage process to study if new information is used to update the initial track advice.

We find that students with tertiary educated parents receive on average a significantly higher initial advice of 1.7 advice levels, which is almost a complete track. Around 55% of this difference is explained by differences in skills between the two groups of students. Differences in non-cognitive skills seem to play only a small, albeit statistically significant role. The procedure of reevaluation and readjustment leads to an overall increase in advice, but it does not decrease the difference in advise by parental education levels.

Differences in other characteristics, such as family background, school and regional characteristics, also explain an additional share of the gap in advice between children by parental education. The unexplained gap decreases from 44.5% to 28.6% once we take additional information into account. This, however, also means that the decision on the track recommendation does not only depend on the students' skills, but also other resources that they might have, the schools they come from, or the place they live in.

The remainder of this paper is organized as follows. Section 2 presents a short overview of the Dutch education system. Section 3 introduces our data and empirical strategy. In Section 4 we present our main results and a range of robustness analyses can be found in Section 5. Finally, in Section 6 we discuss the implications of our findings.

2 Dutch education system

In the Netherlands, formal education typically starts from the age of 4 with two years of pre-school, followed by six years of general primary education.[5] After children finish primary school, around age twelve, they start in one of several tracks in secondary school. The main tracks are pre-vocational secondary education (vmbo), higher general secondary education (havo), or pre-university secondary education (vwo). About half of all children go to vmbo, which lasts four years and prepares for secondary vocational education (mbo). It is subdivided into different levels, each with a particular combination of vocational training and theoretical education. These are vmbo-basic-vocational (vmbo-b), vmbo-senior-vocational (vmbo-k) and vmbo-theoretical-vocational (vmbo-gt).[6] Havo (5 years) and vwo (6 years) are general education programs leading to university. A havo diploma allows students to attend universities of applied sciences (hbo) while a vwo diploma grants access to research universities (wo). Often, schools offer classes with a combination of adjacent tracks in the first one or two years of secondary school.

School advice and final test

By the end of the 6th grade (age twelve) of primary school, but before the first of March, every child gets to hear their teacher's advice[7] about the secondary school track that fits their ability level the best. This school advice is a subjective expert evaluation of the student's abilities based on his/her achievements, attitudes and interests. It reflects the teacher's expectation of the student's future achievement level during secondary education (De Boer et al., 2010). While formulating their recommendation, teachers are expected to carefully assess the child's cognitive and non-cognitive abilities relying on objective and measurable learning performance and following a clear decision-making procedure with no implicit assumptions.[8]

A few months after receiving their school advice, it is mandatory for students to take a test in order to evaluate their language and math skills. This final test always takes place somewhere between 15 April and 15 May. There are several versions

[5] Participating in education is compulsory from age 5.

[6] Vmbo-gt is in fact the advice for the two tracks "gemengde leerweg" and "theoretische leerweg". While they differ in the final exam at the end of lower secondary education, they are taken together as one level in the tracking advice (i.e. at the beginning of lower secondary education).

[7] We loosely use the term teacher's advice to refer to the school advice. The school advice is not entirely decided by one individual. The process involves other members in the primary school among whom is the main teacher.

[8] Inspectie van het Onderwijs, 2018

of the test available for the primary school to choose from, the most popular one being called the Central End Test (CET). The test scores correspond to certain secondary school tracks. The secondary school level inferred from the nationwide test result constitutes what is called the test advice. Combined advices also exist, such as vmbo-b/k or havo/vwo.

Policy change in 2015

In the past, the test advice and the school advice jointly determined the level of secondary education that is suitable for the student. However, since the school year of 2014/2015, the school advice became the decisive element, giving the final test a much less important role in the transition from primary to secondary education.[9] If the test score indicates a higher advice than the initial advice, the school is mandated to reevaluate the student before the final advice is given. They can, but do not have to readjust the initial advice. For students whose test advice is equal or lower than the initial advice, there is no change between initial and final advice.

3 Data and method

3.1 Data

For the analysis in this paper, we use data from the third round of the cohort study COOL of 2013, linked with microdata from Statistics Netherland (CBS). The COOL study surveys students aged 5 to 18 to track educational careers and measure cognitive development, socio-emotional development and social skills. The data is collected using tests and questionnaires administered to teachers, students and their parents. The COOL sample consists of a representative part and an additional part of disadvantaged schools.[10] In this paper we limit the COOL dataset to students who were in 3rd grade (aged 9) and we use variables measuring their cognitive and non-cognitive skills. These data are then merged with microdata comprising the school and the test advice for the same students at age 12, along with a set of background variables. The final sample consists of 5197 students from 311 schools.

[9] Toetsbesluit PO, 2014

[10] The additional sample of disadvantaged schools is chosen based on the school score indicating the social ethnic composition of the student population in a school. Additional schools with a higher score are added to the representative sample.

Initial and final advice

We use two main outcome variables: the initial and the final school advise. These variables each consist of the following nine advice categories of lower secondary education:

1. basic pre-vocational (vmbo-b)
2. basic pre-vocational / senior pre-vocational (vmbo-b/k)
3. senior pre-vocational (vmbo-k)
4. senior pre-vocational / theoretical pre-vocational (vmbo-k/gt)
5. theoretical pre-vocational (vmbo-gt)
6. theoretical pre-vocational / higher general (vmbo-gt/havo)
7. higher general (havo)
8. higher general / pre-university (havo/vwo)
9. pre-university (vwo)

These are the five main tracks in secondary school and their intermediate combinations. When students receive a test score at their final test that corresponds to a higher school level than the initial school advice they received, they are eligible for a reevaluation. This can lead to a readjustment of the school advice. However, not all under-advised students have their school advice readjusted. Thus, we use a second outcome variable, final advice, in order to consider the effects of the readjustment in our analysis. Figure A.1 in the Appendix illustrates the frequency with which each advice level is given as an initial and as a final advice.

Socioeconomic and demographic background

Our data contains a rich set of socioeconomic background variables, including gender, year and month of birth, age of the parents, household income, household composition, parental involvement[11] and whether or not the student needs special care[12]. The central variable in our analyses, parental education, is measured as a binary variable capturing whether at least one of the parents of a student has a tertiary education.[13]

[11] A continuous variable reported by the teacher on the parents' level of involvement in school and in their support for the kid's learning process on a scale of 1 to 5. A score of 5 means very high involvement.

[12] Due to a physical or mental disability or learning problems.

[13] For ease of interpretation, the variable we use in our analyses takes the value of one if none of the parents has a tertiary education.

Table 1: Descriptive statistics

	complete sample	parental education		migration background	
	all	high	not high	no	yes
initial advice	5.624	6.686	4.980	5.809	5.159
	(2.478)	(2.182)	(2.425)	(2.416)	(2.571)
final advice	5.759	6.795	5.132	5.914	5.371
	(2.459)	(2.130)	(2.433)	(2.394)	(2.575)
parents tertiary educated	0.377			0.423	0.264
	(0.485)			(0.494)	(0.441)
no parent with tertiary education	0.623			0.577	0.736
	(0.485)			(0.494)	(0.441)
no migration background	0.715	0.801	0.663		
	(0.451)	(0.399)	(0.473)		
migration background	0.285	0.199	0.337		
	(0.451)	(0.399)	(0.473)		
household income percentile	50.55	64.37	42.17	57.15	33.96
	(27.45)	(25.71)	(24.97)	(25.07)	(26.14)
male	0.495	0.499	0.493	0.499	0.486
	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)
female	0.505	0.501	0.507	0.501	0.514
	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)
non-cognitive skills (std)	0.000	0.122	-0.0739	0.0424	-0.107
	(1.000)	(0.964)	(1.014)	(0.986)	(1.027)
cognitive skills (std)	0.000	0.342	-0.207	0.152	-0.382
	(1.000)	(0.914)	(0.993)	(0.957)	(1.005)
number of observations	5,196	1,961	3,235	3,716	1,480

Note: Standard deviations in brackets.

Migration background is also captured in a binary variable, indicating whether at least one parent of the student was born abroad. Table 1 shows that both the initial and the final advice are strongly correlated with, in particular, parental educational and migration background. It also shows a clear correlation between parental education and migration background.

Cognitive and non-cognitive skills

Our dataset also includes test scores on language and math in grade 3 that are part of the students' tracking system[14] and scores on a non-school cognitive capacities test (nscct) in grade 3. These tests were thus taken 3 years before the school advice. Based

[14]The language tests assess vocabulary, technical reading and reading comprehension.

on earlier findings on rank-order stability of cognitive skills (Borghans et al., 2008), we assume that the (relative) cognitive skills of students remained stable over this time period. The nscct test aims to estimate the student's learning potential. The test consists of five parts: figure composition, exclusion, series of numbers, categories and analogies.

The COOL dataset also includes measures of non-cognitive skills relating to the student's behavior, their relationship with the teacher and other students and their motivation. On a scale of 1 to 5, the teacher evaluated each students' behavior, working attitude, popularity in class and performance relative to real abilities.[15] The teacher also assessed the students on their teacher-student relationship in terms of dependence, conflict and closeness, while the students self-rated their motivation and well-being in the class and with the teacher. We aggregate both cognitive and non-cognitive skills into one respective variable by using the first principal component of a factor analysis. The factor loadings, as well as the means of single items by parental education background can be found in Table A.5 in the Appendix.

Regional and school characteristics

Finally, the dataset contains information on several regional and school characteristics, including the province, the degree of urbanicity (on a 5-point scale), a school score (indicating the socio-ethnic diversity of a school's student population), the school denomination (the religious or ideological vision on which the school is based) and the type of final test the school uses in the last year of primary school.

3.2 Empirical strategy

Our main goal is to study the relation between students' socio-economic background and the school advice they receive. Hereby we consider three models: 1) an unconditional specification where we only include the parental education level and migration background. 2) In the conditional model we control for students' cognitive and non-cognitive skills, in addition to the parental education and migration indicators. 3) In an extended model we also consider other characteristics of students' backgrounds, as well characteristics of the schools they visit and the regions they live in. This last model is likely to underestimate the impact of SES, in that many of the included variables link back to or are a consequence of parental SES. The extended model

[15] The variable performance indicates to what extent the teacher considers the student's performance at school to reflect the student's skills and abilities. A higher score on performance indicates that the student is doing their best.

is therefore more to be understood as an exploration of which observable factors are linked to SES differences in advice, and are potentially included in a teacher's decision making model.

Since we do not know what precisely determines school advice we start by investigating the relation between our variables of interest and the school advice. We use a simple linear model for this, where we add successively more information, estimating the unconditional model, the specification that conditions on skills, and the extended specification. For each type of advice (initial and final) we estimate

$$\text{advice}_i = \alpha + \beta_1 \text{ SES}_i + \beta_2 \text{ migrant}_i + \gamma \text{ skills}_i + \delta Z_i + \epsilon_i,$$

where Z captures the additional background information for student i. This equation relates to the extended model, in the conditional and unconditional model we leave out Z, and Z and the skills, respectively.

To study the difference in advice between groups of students by parental education, we apply an Oaxaca–Blinder decomposition. This allows us to divide the gap into a part that is explained by observable differences between the groups and a part of the gap that remains unexplained. The group of children with high educated parents is denoted as A and the group whose parents are not tertiary educated is denoted as group B.

$$Y_A - Y_B = \bar{X}_A\hat{\beta}_A - \bar{X}_B\hat{\beta}_B = \underbrace{(\bar{X}_A - \bar{X}_B)\hat{\beta}_P}_{\text{explained}} + \underbrace{\bar{X}_A(\hat{\beta}_A - \hat{\beta}_P) + \bar{X}_B(\hat{\beta}_P - \hat{\beta}_B)}_{\text{unexplained}}$$

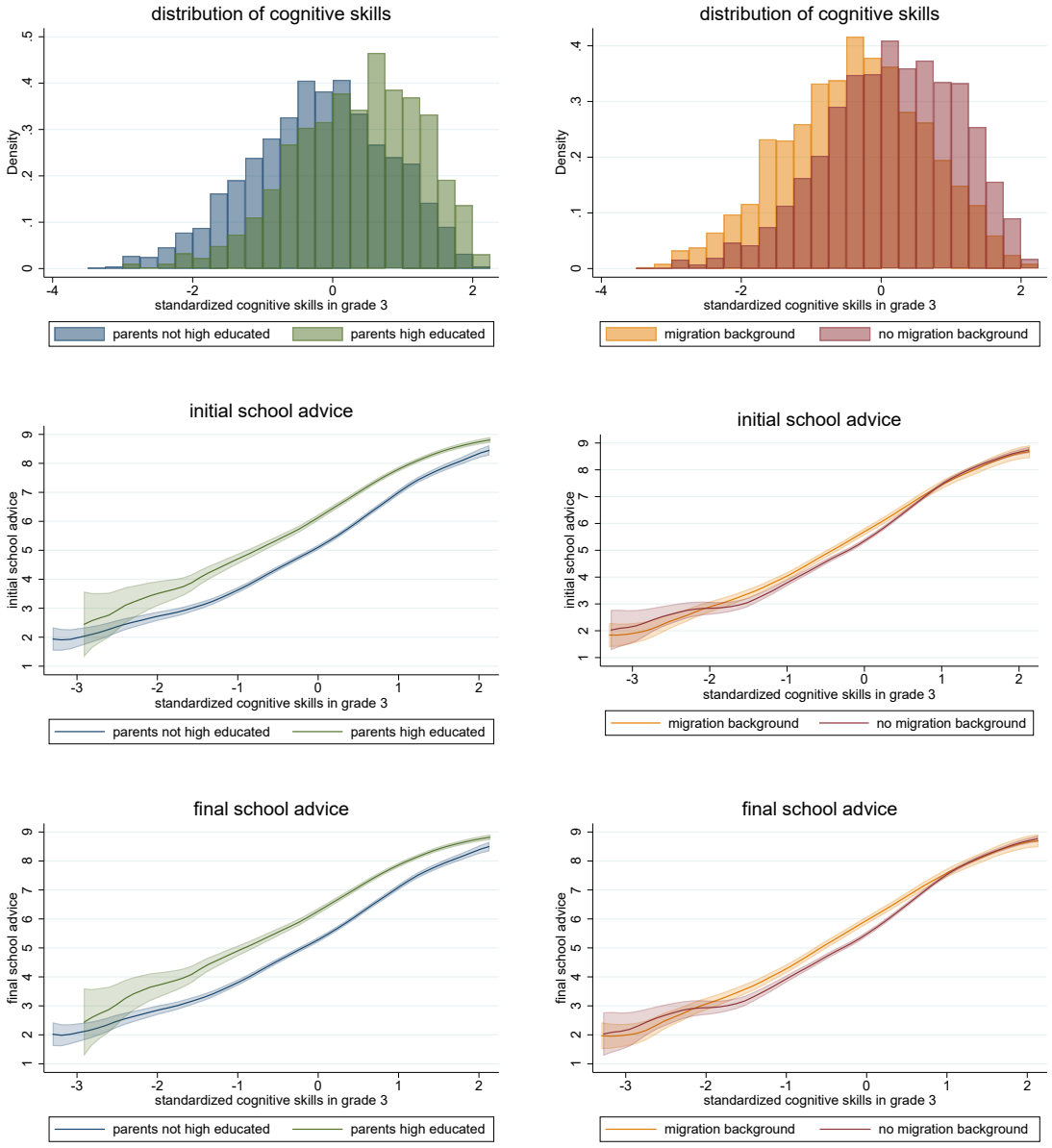
In the conditional model X contains skills measures, while in the extended specification it contains also additional information on the students background, school and region. The subscript P indicates that the estimates result from a pooled estimation including both A and B.

4 Results

Looking at the raw data we start out with a significant unconditional difference in initial advice by parental education of around 1.7 advice levels, out of a total range of 9 possible advice levels. We also find a gap in advice by migration status, however much smaller.

Earlier studies (see e.g. Van Huizen, 2018) find that children from higher SES families have higher cognitive skills throughout and at the end of primary school. We

Figure 1: Distribution of cognitive skills and corresponding school advice



should therefore expect that a significant share of the difference in advice is due to different abilities and prior performance. The top left panel of Figure 1 indeed shows that children from parents with a tertiary education outperform their classmates from

lower educated parents in grade 3, three years before the tracking advice is given. We find a similar pattern for migration background, with migrant children on average performing worse than children without migration background.

The middle panels in Figure 1 display the initial school advice a child receives, conditional on their performance in grade 3. On the left, the local polynomial is split by parental education, on the right by migration background. Unlike with the overall skills distribution, in these panels the patterns that we observe for parental education and migration background clearly differ. We find that almost throughout the skills distribution, children from lower educated parents receive systematically lower advice than children whose parents are tertiary educated. For migration background we do not observe such a gap, but rather that children with a migration background seem to receive a slightly higher advice than their native peers who perform similarly in grade 3. The lowest panels show the relation between cognitive skills and the final advice, that is given after the chance of reevaluation. The overall level of advice increases slightly increases for all groups. While the gap by parental education does not visibly change, we find a slight increase of the positive evaluation of children with migration background in the final advice.

4.1 Initial school advice

To gain a better understanding of the gap in advice by parental education (as shown in the middle panel of Figure 1) we decompose the total difference in advice into an explained part that captures the different characteristics of the two groups, and an unexplained part. However, since the advice-making process is not governed by a strict set of rules, we first have to identify the important input factors.

Determinants of school advice

Table 2 provides the empirical findings on which variables play a role in predicting school advice. The first column captures the unconditional gap, where we only control for parental education and migration background.[16] Column two shows the conditional difference in advice, if we control for a child's skills level. Non-cognitive skills positively relate to a higher advice, however not very strongly. One standard deviation increase in non-cognitive skills relates to a higher advice of about 0.2 levels in initial advice. Teachers may take non-cognitive skills into account especially if there

[16]While these coefficients are unconditional on the students skills levels, the two included variables are correlated. The completely unconditional coefficient would be -1.706*** for parental education, and -0.650*** for migration background respectively.

Table 2: Determinants of initial school advice

VARIABLES	(1) initial advice	(2) initial advice	(3) initial advice	(4) initial advice
parents not high educated	-1.652***	-0.792***	-0.488***	-0.382***
	(0.080)	(0.067)	(0.057)	(0.052)
with migration background	-0.387***	0.385***	0.465***	0.464***
	(0.099)	(0.080)	(0.074)	(0.069)
non-cognitive skills		0.208***	0.139***	0.179***
		(0.029)	(0.028)	(0.026)
cognitive skills		1.708***	1.583***	1.620***
		(0.030)	(0.031)	(0.025)
cognitive skills squared		0.151***	0.156***	0.151***
		(0.018)	(0.018)	(0.017)
background variables	no	no	yes	yes
school fixed effects	no	no	no	yes
constant	yes	yes	yes	yes
Observations	5,196	5,196	5,196	5,196
R-squared	0.116	0.552	0.603	0.668

Note: The additional background variables include demographic variables, family background, school characteristics and regional charactristics. The standard errors are clustered at school level, significant at *** p<0.01, ** p<0.05, * p<0.1

is a doubt on how to place a student, and the student's prior performance does not lead to a clear answer. However, while we measure a broad spectrum of potentially relevant non-cognitive skills (such as work attitude, independence and motivation) we cannot exclude that we are also missing some relevant skills that teachers do observe. As expected, earlier measured cognitive skills relate strongly to the advice that is given. This relation is stronger at the higher end of the skills distribution, where students receive an advice for the more academic tracks. Conditional on the students skills the coefficient on parental education drops to around half the size of the unconditional coefficient, while the coefficient for migration background changes its sign.

There is a broad range of other characteristics of the students themselves, their environment, but also the teachers and schools that could have an impact on the

advice that is given.[17] In Column 3 we include additional information on the child (such as age of the child and gender), their family (e.g. household income, number of siblings or parental involvement), the school (public school versus schools with a religious of philosophical affiliation, socio-economic composition of the students) and the region (degree of urbanicity and province). Table A.1 in the Appendix provides the detailed regression outcomes. We see that controlling for these additional variables further decreases the coefficient on parental education. It is however important to note that this specification is likely to lead to an underestimation of the coefficients on parental education, due to adding factors that are closely related to the parents education level, such as household income, but also region or the type of the chosen primary school.[18] Finally, there could also be differences in advising between schools, i.e. the guidelines on how to form an advice could differ between schools, or average teacher characteristics.[19] To account for these potential between-school differences in the advice making process we add school fixed effects in Column 4. While the coefficient on parental education further decreases, once we take school fixed effects into account, it remains sizable and statistically significant.

Decomposition of the advice gap

The group of children who have at least one parent with a tertiary education receives an average advice of 6.7, which is between a mixed vmbo gt–havo and a havo advice. A havo diploma is the first diploma that gives direct access to higher education. 60% of the students from high educated parents receive an initial advice for at least a havo track. The group of students whose parents are not higher educated receives a lower average advice of 5, vmbo gt, which is the highest track in vocational education. Students who follow this track can transition into any track of post secondary vocational education. The highest track of post secondary education in turn also provides access to higher education. The total raw difference of 1.7 is

[17] For example, Verhagen (2021) shows that differences between schools are important predictors for the gap in advice.

[18] The Netherlands has a system of free school choice, where parents can enroll their children at any school of their preference. This leads to a partial segregation as Oosterbeek et al. (2021) show for the case of Amsterdam.

[19] Teacher bias can be linked to teacher quality. Lavy and Megalokonomou (2019) find a stronger gender bias among low value added teachers. This is mainly a problem if certain schools systematically attract different teachers. The Netherlands currently faces a shortage of teachers in primary education, and this mostly manifests in schools with a larger share of low SES students (https://www.onderwijsinspectie.nl/onderwerpen/staat-van-het-onderwijs/trends-in-het-onderwijs/leraren-en-lerarentekort/oplopend-lerarentekort-bedreiging-voor-gelijke-kansen-in-het-onderwijs).

Table 3: Decomposition of difference by parental education in initial school advice

	(1) initial advice		(2) initial advice	
mean by parental edu:				
high educated	6.686		6.686	
not high educated	4.980		4.980	
difference	1.706		1.706	
total explained	0.946***	55.5%	1.217***	71.3%
	(0.057)		(0.068)	
total unexplained	0.759***	44.5%	0.488***	28.6%
	(0.068)		(0.058)	
explained:				
non-cognitive skills	0.041***	2.4%	0.027***	1.6%
	(0.009)		(0.007)	
cognitive skills	0.917***	53.8%	0.869***	50.9%
	(0.056)		(0.053)	
cognitive skills squared	-0.012*	-0.7%	-0.012*	-0.7%
	(0.006)		(0.007)	
background			0.261***	15.3%
			(0.033)	
school			0.077**	4.5%
			(0.030)	
regional			-0.005	-0.3%
			(0.017)	
Observations	5,196		5,196	

Note: Oaxaca–Blinder decomposition. The additional background variables include demographic variables, family background, school characteristics and regional characteristics. The standard errors are clustered at school level, significant at *** p<0.01, ** p<0.05, * p<0.1

almost a full advice level (2 levels in mixed advice), and relates to later educational opportunities, since it determines the access to further education.

In a system where academic performance is the key input factor for the tracking advice, we should expect that most of the difference in advice would be explained by

previous skills measures. However, in Table 3 we see that only 55% of the difference in advice is explained by observable differences in cognitive and non-cognitive skills between the two groups. Zooming in, we see that it is mainly differences in cognitive skills that explain the gap. Differences in non-cognitive skills are also statistically significant, but explain only 2.4% of the gap, or a level difference of 0.041, while 45% of the gap remains unexplained. The unexplained difference could partly be related to "teacher bias", i.e. that students from different groups are rewarded differently for specific skills. It can also contain some additional differences in characteristics that we do not observe, but teachers do. However, given the broad range of tests and non-cognitive skills measures it is unlikely that teachers hold additional information on skills that explains the remaining 45% of the gap.

If we relax the notion that only skills have an impact on the advice, and allow for a model in which teachers also consider other factors, such as parental support (e.g. through the option for paying for tutoring) or whether all types of tracks are readily available in the neighborhood (as captured by the degree of urbanicity), we are able to explain a larger share of the gap. We find that differences in additional individual and family background, as well as in school and regional characteristics explain together with skills up to 71% of the unconditional gap. However from a perspective of equality of opportunities, a fraction of the now explained differences can also be considered as a bias. For example the 7% of the difference in advice[20] that is explained by differences in household income might be based on anticipated access to private tutoring, but it also decreases the equality of opportunities for students from low income families beyond the impact it has on their performance within a given track. Another share of the difference is explained by the unequal probability of attending a school with many students from low SES families, which is related to lower advices. Regional differences do not provide additional explanation in our sample.

Since gender is equally distributed within both groups, gender does not explain any of the difference in the advice between the two groups. It could however be that the gap differs by gender and that it is explained by different characteristics (Cornwell et al., 2013). Both in terms of average advice and difference by parental education, girls and boys are very similar. Also the share of the gap that is explained by skills is very comparable for girls and boys. We do see however that additional characteristics explain more of the gap for boys, in particular differences in background characteristics explain a larger part (a total of 18% for boys and 12% for girls respectively).[21]

[20] See Table A.2 in the Appendix for results on the individual control variables.

[21] Tables A.3 and A.4 in the Appendix provide the regression output for analyses by gender.

4.2 Adjustments to final school advice

The initial advice is important, and for most students it equals the final advice and determines into which track of secondary education they are placed. However, students do have a chance to improve on the initial advice. All students have to take a final exam, and if in this exam they outperform their initial advice, the schools have to reevaluate their advice. Approximately a third of these reevaluations lead to an upwards adjustment, and thus to a higher final advice (Swart et al., 2019). The process of reevaluation is an opportunity for previously under-advised students to receive a track advice that matches their ability. Since students from lower educated parents are receiving lower advices given their previous performance, we could expect that they are more often reevaluated and readjusted. The gap in the final advice should therefore be smaller.

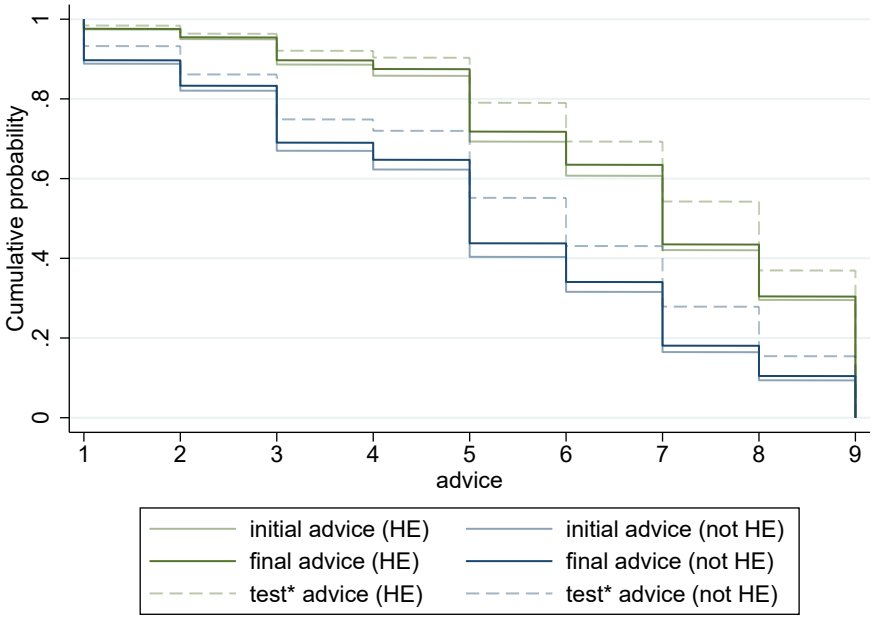
However the gap by parental education seems to stay the same for the final advice as it is for the initial advice. Figure 2 shows the cumulative distribution functions for both advices. Since the initial advice can only be upwards adjusted, but not downwards, the curves will shift to the right if at least a fraction of the students is reevaluated and receives a readjustment. We indeed see a shift to the right for both groups, moreover, this shift seems of comparable size for both groups of children. This indicates that the process of reevaluation does not lead to catching up for the lower advised group.

To test if the change in the conditional impact of parental education on the advice is statistically significant we jointly estimate the linear regression of advice on skills and parental education for both types of advice (Columns 2 and 5 in Table A.1 in the Appendix). The coefficient on parental education changes only marginal and we cannot reject the null hypothesis that they are in fact of the same size.

The persistence of the gap in the final advice begs the question of why the procedure does not decrease the gap, while it is designed to allow under-advised students to get an advice that is closer to their actual ability. Any difference between initial and final advice stems from either the test performance, relative to initial advice, or the school's decision to grant a readjustment. We can shed some light on this issue by looking at a hypothetical final advice, in which the test advice always leads to an upwards adjustment. It is important to stress that the test advice only has an impact if it is higher than the initial advice. This leads to an incentive structure where students with a satisfactory advice are less incentivized to perform well at the test. We can therefore not use the test outcome directly, but have to take the highest advice among the test and initial advice.

The dashed line in Figure 2 indicates the advice students would have gotten, if the test advice was followed in all instances when it surpassed the initial advice

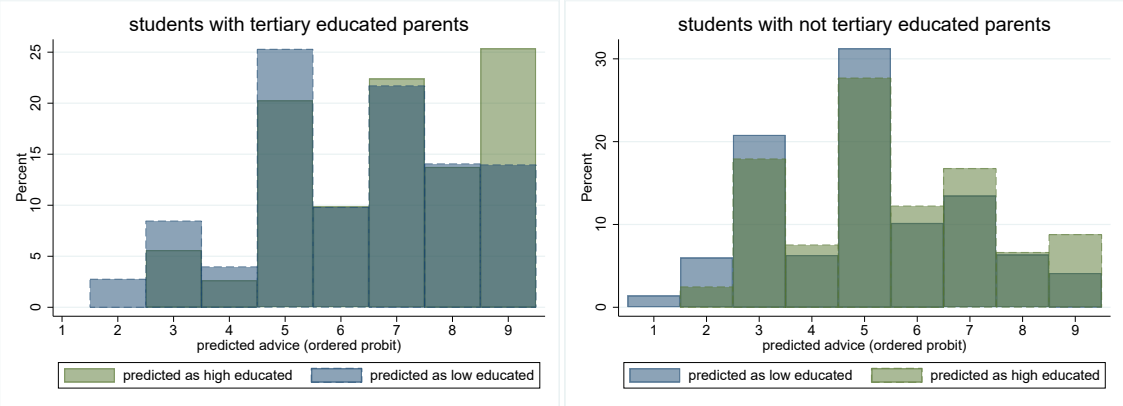
Figure 2: Initial versus final advice by parental education



(test*). We see that in particular on the lower advice levels, the gap by parental education would decrease. The joint estimation and testing of the coefficient of parental education (conditional on skills) on initial and test* advice supports this finding. The decrease of the coefficient, from -0.79 for the initial advice to -0.61 for the test* advice, is statistically significant. This is in line with the findings of an earlier study that, given a right for reevaluation, students from lower educated parents are less likely to receive a readjustment (Swart et al., 2019). However, there has also been recent critique to using the test advice to analyze under-advice, due to concerns of biased measurement error (Van Huizen, 2021).

So while we cannot pinpoint the causal impact of parental education on advice, we see that, conditional on having the same previously measured skills, there is a significant and persistent gap in both initial and final advice between children from different socio-economic backgrounds.

Figure 3: Non-linear decomposition



5 Robustness analyses

5.1 Non-linear decomposition to account for ordinal outcome variable

In our main analyses we treat the outcome variable as continuous. However, not all advice levels are equally likely to be given, in particular mixed advices are less frequent. Hence, a linear approach might be biased since it assumes equal distance between all levels. We relax this assumption and estimate a non-linear decomposition. Following Verhagen (2021) we use an ordered probit model to estimate the effect of skills on advice separately by parental education, and predict the conditional advice for both groups twice, once using the estimates from the respective group, and once using the estimates from the other group.[22]

The left panel in Figure 3 shows the predicted advice that students with tertiary educated parents would get, based on the estimation from the group with high educated parents and the group with not high educated parents. The right panel shows the same for the sample of students whose parents are not tertiary educated. We see that also in this specification students from higher educated parents are more likely to receive a higher advice, while with the same characteristics they would receive a significantly lower advice if they were treated as students from lower educated parents. And the same holds for students from lower educated parents, they would receive a higher advice if they were treated as students from high educated parents.

[22]We split the sample into an estimation and a prediction sample. Our predictions are thus based on a subgroup of our normal sample that has not been used to estimate the prediction parameters.

The share of the total difference that is explained by skills in this model is comparable to what we find in the linear model.

Alternatively, we can also apply a simplified binary outcome variable that captures if a student receives an advice for general education (havo and above), which will later grant them direct access to higher education. Looking at the unconditional difference we see that 60% of all students with high educated parents receive an advice for general education, while only 32% of student with lower educated parents do so. We run the same decomposition on these differences as we do for our main analyses. The result is very similar, differences in skills explain around 57% of the difference in receiving an advice for general education.[23]

5.2 Sensitivity to aggregation of skills measures

We are able to measure a broad range of skills, but are relatively agnostic about how important a specific skill is for future success, or the teacher's prediction thereof. In our main analyses we combine all cognitive skills and all non-cognitive skills into one respective aggregate variable by running a factor analysis and taking the principal component. If teachers weigh specific skills higher than the average overall skills, we might throw away information by using only a combined measure. To test how large the additional variance is that could be explained by allowing for different weights on skills we run the main analyses again but now include all skills items individually. While this approach may lead to a slight over-fitting of our model, it also provides an upper bound on what we could potentially explain with the skills we observe. In the OLS regressions, the explained variance (in a model where we only control for skills and parental education) grows from 0.54 with combined skills measures to 0.63 when adding the skills individually. If we now investigate how much more of the gap in advice could be explained by differences in skills if added individually, we find only a very small change. The explained part of the gap is 57% in this more flexible model, as opposed to 55.5% in the more parsimonious approach.

5.3 Sensitivity to definition of parental education

The gap in advice by parental education captures the difference between the group of children with tertiary educated parents (37.7%) versus all children whose parents do not have a university degree, regardless of whether they have a professional education or are low educated. In this sensitivity analysis we test whether the difference is specifically relevant for a high parental educational level or if there is also

[23]The regression output for the robustness checks is in the Webappendix.

a significant difference between the groups of children whose parents have a middle or low educational level.

We start out by including separate categories in the linear regression on advice, instead of the binary variable that captures whether parents have a tertiary education or not. We find that there is also a difference between the low and middle level of parental education, with low educated parents as a base group, the coefficients in the conditional specification (controlling for skills) are 0.339*** for a middle level and 0.867*** for a high level of parental education, respectively. To decompose the gap we need to split the sample into two groups. We hence define a binary variable that captures low educated parents versus middle and high educated parents. Around 20% of the students in our sample fall into the category of having low educated parents.

Students with low educated parents receive on average an advice of 4.3, as opposed to an average advice of 5.9 that the group of children with middle and high educated parents gets. The unconditional gap in advice of 1.66 is comparable to the gap that we observe in our main analysis. However, skills explain a significantly larger part of this gap, with an explained part of 70.6% by skills alone, and 84.4% if we also include other characteristics.

5.4 Representativeness

Population weights

The COOL cohort study on which our sample is based, over-samples schools with a higher share of students with migration background and from low SES families. The average advice that students receive in our sample is lower compared to the national average. This holds also within groups. While in 2016 the national average initial school advice for children with high educated parents was 7.02, in our sample it is 6.69. Among the children with not high educated parents the national average was 5.23 compared to our sample average of 4.98.

We construct population weights for our sample on the basis of the complete 2016 cohort. Using those population weights we find an even larger unconditional gap (1.776) as well as a stronger conditional relation between advice and parental education (-0.888***). The difference of the gap that is explained by skills amounts to 50.8%. These results indicate that our main results might be slightly underestimating the role of parental education in the school advice, while slightly overestimating the importance in differences in skills to explain this gap. However, our population weights are not very precise; the reweighed group averages in advice come closer to the averages within the national cohort, but do not match exactly (with a reweighed average of 6.8 for high educated and 5.02 for not high educated).

These deviations could be partially due to less precision in the information on parental education in the sample of the complete cohort. While we know the parental education for almost all children in our sample, the group for whom the education is not known in the national sample is about a quarter. Since our binary variable captures (known) high educated, we include the children with missing information in the group with middle and low educated parents. This could mean some of the children in this sample do have in fact higher educated parents, and the gap that we observe in the national sample is an underestimation.

Transition effects due to new system

The cohort we study was only the second cohort after the introduction of a new regulation on how the final school advice is formed. To be able to tell more about whether our results are also likely to hold for newer cohorts, or whether we measure mostly transitory effects, we studied trends in average advice for different subgroups. Figure A.2 in the Appendix illustrates that the average advice by parental education level, as well as by migration background has been stable over time, and that we neither observe a closing of the gap over time nor an increase in difference.

6 Discussion

In this study we investigate the significant difference in school track advice children from different socio-economic backgrounds receive. Unconditional on skills and other characteristics, the group of students with high educated parents receives a higher advice of 1.7 levels, out of 9 possible levels. Our main analysis estimates that approximately 55% of this gap by parental education can be explained by differences in skills. Additional robustness analyses confirm this finding, the explained part is between 50% and 60% in a range of different specifications of the estimation model, definition of the dependent variable, the aggregation of skills measures or using population weights.

The unexplained difference of 40 to 50% is still a significant gap in advice that can potentially harm the equality of opportunities for children whose parents are not high educated. There are several other factors which can potentially add to explaining the remaining gap, but are unlikely to fully explain it. As we discuss in Section 5.2, we might not observe all relevant skills that teachers take into account. Among the cognitive skills we do however not only measure math and language — which are unarguable key, and also the ones that are tested in the central exam — but we also include measures for the ability to grasp concepts, which is important for future

learning. The non-cognitive skills are also measured very broadly, with concepts such as motivation, work attitude and independence. While here we can be less certain that we measure all of the most important concepts, the ones that we do measure have little impact, and it is hard to imagine that a specific other trait would explain much more.

The range of different skills that are tested, spaced over time, also partially alleviates a potential second concern, namely measurement error. If measured with noise, our estimates of the impact of skills will be attenuated. Van Huizen (2021) discusses the relation between measurement error and under-advice, using the nationwide test outcomes as skills measures, which are elicited in one exam. Our skills measures are much less sensitive to this type of measurement error (i.e. effects of a lucky day), since they are measured over multiple test days and span over a wider range of skills. In contrast, we could encounter a different type of measurement error through our early measure of skills. We use grade three skills as a proxy for skills in grade six. Random events that occur between grade three and six could thus introduce noise to our skills proxy. However, cognitive skills already reach high of rank-order stability by age nine (Borghans et al., 2008), such that it is unlikely that measurement error would explain the complete difference in advice between groups or even a substantial part of it.

Differential growth of skills between the two groups could yet be another potential candidate that would lead to an underestimation of the explanatory power of skills. Van Huizen (2018) studies the development of skills, using earlier cohorts of the COOL cohort study.[24] He finds that most of the achievement gap by parental education can be attributed to differences that are already measurable in kindergarten. However, there is a slight increase of the achievement gap during primary school. Between grades three and six he finds that there is a slight increase in the group differences in math and language skills of around 10% between children from high educated parents versus children whose parents are not high educated. If we take a diverging development of skills of 10% into account in our analyses, we could expect to explain a bit more of the school advice gap, around 60%, which still leaves a large part unexplained.

We cannot exclude that we underestimate the share of the gap that is explained by skills. However, measurement error and unobserved skills are unlikely to explain the remaining 44.5% of the total difference, which amounts to 0.76 advice levels. This leaves room for other factors, that are not skills-related, to also matter in the process of advice making, and in the difference in advice that we observe between children

[24]The COOL cohort study runs until 2013. We can therefore not observe students' COOL performance in grade six under the new system of school-advice, which starts in 2015.

from different backgrounds.

Our additional analyses show that there are indeed other factors that are predictive of school advice and can explain an additional share of the gap. This can be either through teachers processing other information which is not based on a students' skills into their assessment, or through differences in advising between schools or individual teachers. Many of those additional factors are directly related to parental education and might be more mediators through which parental education is related with advice, rather than independent effects. And taking for example additional resources into account for the school track advice, such as financial resources to pay for private tutoring can additionally increase the negative effect that early tracking has on the equality of opportunity. However, even after taking the additional information on individual and family background, school and regional characteristics into account, a difference of 0.49 advice levels (28.6% of the total gap) remains. This still unexplained difference could be in part due to (implicit) differential treatment of students on the basis of their socio-economic background, also referred to as teacher bias.

Several possible options for reducing this gap have been discussed. Falk et al. (2020) show that a low-intensity mentoring intervention can decrease the gap. Verhagen (2021) points to the importance of between school differences in advising, while Alesina et al. (2018) show that informing teachers about implicit bias leads to an adjustment of behaviour. Very little information is available on the fit of the students within their new tracks in secondary education. The overarching organization of Dutch primary and secondary schools has recently suggested that more systematic feedback from the receiving schools is needed to make better school advices. This could also help to point out potential teacher bias, although more research on this topic is needed. Finally, the Onderwijsraad, a Dutch education think tank, recommended to altogether delay tracking by three years.[25]

[25] https://www.onderwijsraad.nl/publicaties/adviezen/2021/04/15/later-selecteren-beter-differentieren

References

Alesina, A., Carlana, M., La Ferrara, E., and Pinotti, P. (2018). Revealing stereotypes: Evidence from immigrants in schools. NBER working paper No. 25333.

Borghans, L., Diris, R., Smits, W., and De Vries, J. (2019). The long-run effects of secondary school track assignment. *PloS one*, 14(10):e0215493.

Borghans, L., Diris, R., Smits, W., and De Vries, J. (2020). Should we sort it out later? The effect of tracking age on long-run outcomes. *Economics of Education Review*, 75:101973.

Borghans, L., Duckworth, A. L., Heckman, J. J., and Ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources*, 43(4):972–1059.

Canaan, S. (2020). The long-run effects of reducing early school tracking. *Journal of Public Economics*, 187:104206.

Card, D. and Giuliano, L. (2016). Can tracking raise the test scores of high-ability minority students? *American Economic Review*, 106(10):2783–2816.

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. *The Quarterly Journal of Economics*, 134(3):1163–1224.

Cornwell, C., Mustard, D. B., and Van Parys, J. (2013). Noncognitive skills and the gender disparities in test scores and teacher assessments: Evidence from primary school. *Journal of Human Resources*, 48(1):236–264.

De Boer, H., Bosker, R. J., and Van der Werf, M. P. (2010). Sustainability of teacher expectation bias effects on long-term student performance. *Journal of Educational Psychology*, 102(1):168.

Falk, A., Kosse, F., and Pinger, P. (2020). Mentoring and schooling decisions: Causal evidence. IZA discussion paper No. 13387.

Geven, S., Batruch, A., Van de Werfhorst, H., et al. (2018). Inequality in teacher judgements, expectations and track recommendations: A review study. *Amsterdam: Universiteit van Amsterdam.*

Hanna, R. N. and Linden, L. L. (2012). Discrimination in grading. *American Economic Journal: Economic Policy*, 4(4):146–68.

Korthals, R., Schils, T., and Borghans, L. (2021). Track placement and the development of cognitive and non-cognitive skills. *Education Economics*, pages 1–20.

Lavy, V. and Megalokonomou, R. (2019). Persistency in teachers' grading bias and effects on longer-term outcomes: University admissions exams and choice of field of study. NBER working paper No. 26021.

Oosterbeek, H., Sóvágó, S., and Van der Klaauw, B. (2021). Preference heterogeneity and school segregation. *Journal of Public Economics*, 197:104400.

Papageorge, N. W., Gershenson, S., and Kang, K. M. (2020). Teacher expectations matter. *Review of Economics and Statistics*, 102(2):234–251.

Pekkala Kerr, S., Pekkarinen, T., and Uusitalo, R. (2013). School tracking and development of cognitive skills. *Journal of Labor Economics*, 31(3):577–602.

Roller, M. and Steinberg, D. (2020). The distributional effects of early school stratification—non-parametric evidence from Germany. *European Economic Review*, 125:103422.

Sprietsma, M. (2013). Discrimination in grading: Experimental evidence from primary school teachers. *Empirical Economics*, 45(1):523–538.

Sulzmaier, D. (2020). The causal effect of early tracking in German schools on the intergenerational transmission of education. BGPE discussion paper No. 187.

Swart, L., Van den Berge, W., and Visser, D. (2019). De waarde van eindtoetsen in het primair onderwijs. CPB notitie.

Timmermans, A. C., Kuyper, H., and Van der Werf, G. (2015). Accurate, inaccurate, or biased teacher expectations: Do Dutch teachers differ in their expectations at the end of primary education? *British Journal of Educational Psychology*, 85(4):459–478.

Van Elk, R., Van der Steeg, M., and Webbink, D. (2011). Does the timing of tracking affect higher education completion? *Economics of Education Review*, 30(5):1009–1021.

Van Huizen, T. (2018). The evolution of achievement gaps from early childhood to adolescence in the Netherlands. In Passaretta, G. and Skopek, J., editors, *Roots and Development of Achievement Gaps - A Longitudinal Assessment in Selected European Countries*, page 50. ISOTIS.

Van Huizen, T. (2021). Teacher bias or measurement error bias? U.S.E working paper No. 21-13.

Verhagen, M. D. (2021). Prediction in educational research: An application to the study of teacher bias. SocArXiv No. y6mnb.

A Appendix

Table A.1: Determinants of school advice

VARIABLES	(1) initial advice	(2) initial advice	(3) initial advice	(4) final advice	(5) final advice	(6) final advice
non-cognitive skills		0.208***	0.139***		0.205***	0.139***
		(0.029)	(0.028)		(0.029)	(0.029)
cognitive skills		1.708***	1.583***		1.682***	1.559***
		(0.030)	(0.031)		(0.030)	(0.031)
cognitive skills squared		0.151***	0.156***		0.122***	0.128***
		(0.018)	(0.018)		(0.018)	(0.018)
parents not tertiary edu	-1.652***	-0.792***	-0.488***	-1.624***	-0.776***	-0.482***
	(0.080)	(0.067)	(0.057)	(0.079)	(0.066)	(0.058)
with migration background	-0.387***	0.385***	0.465***	-0.285***	0.482***	0.521***
	(0.099)	(0.080)	(0.074)	(0.097)	(0.081)	(0.074)
household income percentile			0.005***			0.005***
			(0.001)			(0.001)
gender			-0.175***			-0.162***
			(0.048)			(0.047)
type household			0.007			0.020
			(0.069)			(0.069)
special needs student			-0.594***			-0.596***
			(0.065)			(0.065)
age student (in month)			-0.058***			-0.061***
			(0.004)			(0.004)
age mother			0.007			0.004
			(0.006)			(0.006)
age father			0.007			0.008
			(0.005)			(0.005)
number of siblings			0.050**			0.043*
			(0.024)			(0.024)
parental involvement			0.175***			0.166***
			(0.037)			(0.036)
school SES high			-0.114			-0.077
			(0.096)			(0.091)
school SES middle			-0.077			-0.110
			(0.124)			(0.116)
school SES low			-0.146			-0.187
			(0.149)			(0.145)
school SES very low			-0.589***			-0.651***
			(0.191)			(0.198)
final exam Route8			-0.214*			-0.211*
			(0.121)			(0.120)
final exam IEP			-0.056			-0.047
			(0.091)			(0.091)
final exam other			-0.159			-0.227
			(0.373)			(0.358)
school protestant christian			0.108			-0.273
			(0.196)			(0.197)
school roman catholic			-0.046			-0.404**
			(0.192)			(0.195)

school other christian			0.029			-0.320*
			(0.190)			(0.194)
school other religion			0.515			0.140
			(0.319)			(0.396)
school special philosophy			0.379			-0.009
			(0.253)			(0.259)
urbanicity high			0.005			0.007
			(0.125)			(0.124)
urbanicity medium			-0.187			-0.199
			(0.138)			(0.136)
urbanicity low			-0.190			-0.169
			(0.131)			(0.130)
urbanicity very low			-0.313**			-0.340**
			(0.155)			(0.153)
region fixed effect			yes			yes
Constant	6.763***	5.857***	12.872***	6.852***	5.983***	13.796***
	(0.073)	(0.076)	(0.719)	(0.070)	(0.073)	(0.722)
Observations	5,196	5,196	5,196	5,196	5,196	5,196
R-squared	0.116	0.552	0.603	0.110	0.542	0.596

Note: Several variables are categorical with the following categories as a base group: "very high" for school SES, "CITO" for final exam, "public school" for school type, and "very high" for degree of urbanicity. We further include binary variables for missings in fathers' age and missing measurement of parental involvement. The standard errors are clustered at school level, significant at *** p<0.01, ** p<0.05, * p<0.1

Table A.2: Decomposition of difference by parental education in school advice

VARIABLES	(1) initial advice		(2) final advice		(3) final advice	
mean by parental edu:						
high educated	6.686		6.795		6.795	
not high educated	4.98		5.132		5.132	
difference	1.706		1.663		1.663	
total explained	1.217***	71.3%	1.182***	71.1%	1.182***	71.1%
	(0.068)		(0.066)		(0.066)	
total unexplained	0.488***	28.6%	0.482***	29.0%	0.482***	29.0%
	(0.058)		(0.058)		(0.058)	
explained:						
non-cognitive skills	0.027***	1.6%	0.027***	1.6%	0.027***	1.6%
	(0.007)		(0.007)		(0.007)	
cognitive skills	0.869***	50.9%	0.856***	51.5%	0.856***	51.5%
	(0.053)		(0.053)		(0.053)	
cognitive skills squared	-0.012*	-0.7%	-0.010*	-0.6%	-0.010*	-0.6%
	(0.007)		(0.005)		(0.005)	
background			0.248***	14.9%		
			(0.033)			
school			0.073**	4.4%		
			(0.031)			
regional			-0.012	-0.7%		
			(0.017)			
with migration background	-0.064***	-3.8%			-0.072***	-4.3%
	(0.014)				(0.014)	
household income percentile	0.118***	6.9%			0.114***	6.9%
	(0.026)				(0.027)	
age student (in month)	0.097***	5.7%			0.102***	6.1%
	(0.011)				(0.012)	
parental involvement	0.077***	4.5%			0.073***	4.4%
	(0.017)				(0.016)	
school SES very low	0.047***	2.8%			0.052***	3.1%
	(0.018)				(0.019)	
Observations	5,196	5,196	5,196	5,196	5,196	5,196

Note: Oaxaca–Blinder decomposition. All background variables from Table A.1 are included in Column 1 and 3. Only variables that are statistically significant at least in one specification are reported. The standard errors are clustered at school level, significant at *** p<0.01, ** p<0.05, * p<0.1

Table A.3: Determinants of initial school advice by gender

	boys			girls		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	advice	advice	advice	advice	advice	advice
parents not tertiary educated	-1.654***	-0.768***	-0.430***	-1.650***	-0.812***	-0.553***
	(0.115)	(0.091)	(0.081)	(0.097)	(0.076)	(0.071)
with migration background	-0.372***	0.347***	0.533***	-0.402***	0.432***	0.414***
	(0.140)	(0.102)	(0.100)	(0.115)	(0.094)	(0.095)
non-cognitive skills		0.243***	0.145***		0.207***	0.140***
		(0.038)	(0.039)		(0.044)	(0.042)
cognitive skills		1.677***	1.562***		1.736***	1.618***
		(0.038)	(0.039)		(0.039)	(0.041)
cognitive skills squared		0.180***	0.187***		0.116***	0.124***
		(0.025)	(0.024)		(0.025)	(0.025)
background variables			yes			yes
constant	yes	yes	yes	yes	yes	yes
Observations	2,573	2,573	2,573	2,623	2,623	2,623
R-squared	0.113	0.536	0.597	0.119	0.571	0.617

Note: The additional background variables include demographic variables, family background, school characteristics and regional charactristics. The standard errors are clustered at school level, significant at *** p<0.01, ** p<0.05, * p<0.1

Table A.4: Decomposition of of advice gap advice by gender

	boys		**girls**	
	(1)	(2)	(3)	(4)
	inital advice	inital advice	inital advice	inital advice
mean by parental edu:				
high educated	6.688	6.688	6.683	6.683
	(0.094)	(0.087)	(0.087)	(0.086)
not high educated	4.997	4.997	4.964	4.964
	(0.082)	(0.075)	(0.082)	(0.075)
difference	1.692	1.692	1.719	1.719
	(0.111)	(0.104)	(0.101)	(0.100)
explained	0.942***	1.262***	0.958***	1.166***
	(0.074)	(0.086)	(0.075)	(0.088)
unexplained	0.750***	0.430***	0.761***	0.553***
	(0.094)	(0.080)	(0.077)	(0.071)
explained:				
non-cognitive skills	0.054***	0.032***	0.036***	0.025***
	(0.014)	(0.011)	(0.012)	(0.009)
cognitive skills	0.901***	0.856***	0.931***	0.890***
	(0.071)	(0.068)	(0.074)	(0.072)
cognitive skills squared	-0.014	-0.014	-0.010	-0.010
	(0.011)	(0.011)	(0.007)	(0.007)
background		0.309***		0.210***
		(0.046)		(0.042)
school		0.058		0.087**
		(0.037)		(0.036)
regional		0.021		-0.035
		(0.021)		(0.024)
Observations	2,573	2,573	2,623	2,623

Note: Oaxaca-Blinder decomposition. The additional background variables include demographic variables, family background, school characteristics and regional charactristics. The standard errors are clustered at school level, significant at *** p<0.01, ** p<0.05, * p<0.1

Table A.5: Descriptive statistics, cognitive and non-cognitive skills

	All			parental education		migration background	
				high	not high	native	migrant
	mean	sd	factor loading	mean	mean	mean	mean
cognitive skills							
non-school-related (NSCCT)							
figure comosition	10.30	2.77	0.32	10.81	10.00	10.54	9.70
exclusions	12.01	2.50	0.32	12.44	11.75	12.21	11.50
number series	11.09	2.45	0.32	11.55	10.81	11.24	10.72
categories	15.80	3.22	0.33	16.43	15.42	16.16	14.92
analogies	13.36	3.88	0.36	14.31	12.79	13.70	12.53
school related							
reading comprehension	51.93	28.26	0.38	60.01	47.03	55.26	43.58
technical reading	50.83	28.52	0.16	53.14	49.43	49.60	53.93
vocabulary	51.83	27.92	0.39	59.45	47.20	55.06	43.71
math	52.39	28.15	0.37	60.59	47.41	57.85	38.66
non-cognitive skills							
reported by student							
self-confidence	3.83	0.69	0.16	3.82	3.84	3.78	3.97
motivation	4.16	0.66	0.18	4.12	4.18	4.10	4.31
wellbeing with teacher	3.80	0.61	0.21	3.77	3.82	3.79	3.82
wellbeing with peers	4.10	0.66	0.22	4.09	4.11	4.12	4.05
assessed by teacher							
perfomance (rel)	3.47	0.80	0.30	3.53	3.43	3.51	3.36
behavior	3.74	0.83	0.40	3.83	3.68	3.78	3.63
work attitude	3.46	0.90	0.38	3.59	3.39	3.48	3.43
popularity	3.68	0.76	0.35	3.75	3.64	3.71	3.61
independence	3.97	0.80	0.29	4.08	3.91	4.00	3.90
no conflict	4.34	0.77	0.42	4.43	4.28	4.38	4.23
closeness	3.73	0.63	0.26	3.76	3.71	3.76	3.64
Number of observations	5196			1961	3235	3716	1480

Note: Detailed information on the cognitive tests and specific non-cognitive skills items can be found in the technical report of the COOL cohort study (wave 3, primary education): https://repository.ubn.ru.nl/handle/2066/211243.

Figure A.1: Initial and final school advice

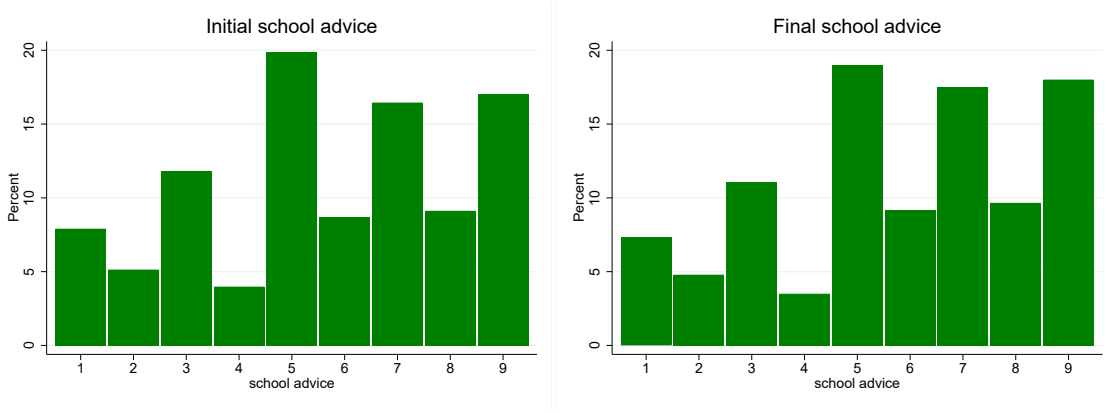
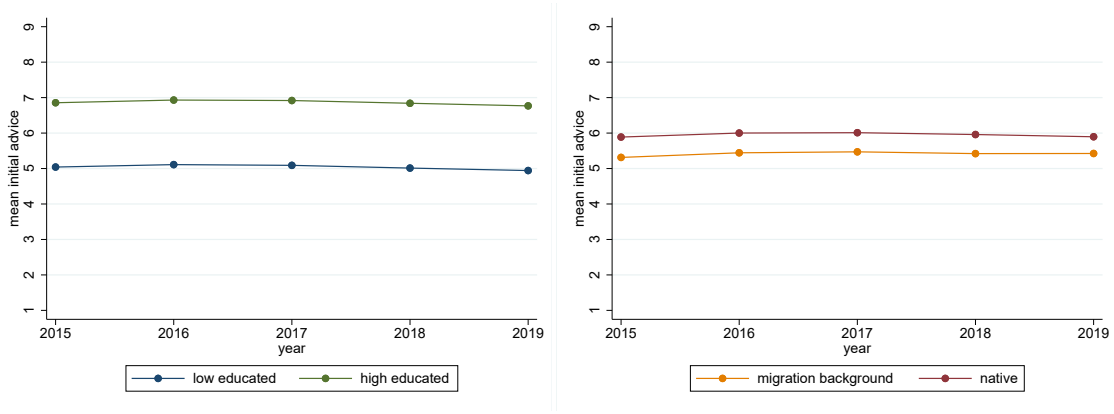


Figure A.2: Trends in school advice over school years



B Webappendix

Table B.1: Determinants of biniary school advice (vocational versus academic)

VARIABLES	(1) initial advice	(2) initial advice	(3) initial advice	(4) final advice	(5) final advice	(6) final advice
parents not tertiary educated	-0.291***	-0.126***	-0.083***	-0.294***	-0.130***	-0.087***
	(0.016)	(0.015)	(0.014)	(0.016)	(0.015)	(0.014)
non-cognitive skills		0.041***	0.034***		0.042***	0.034***
		(0.006)	(0.006)		(0.006)	(0.006)
cognitive skills		0.296***	0.287***		0.292***	0.283***
		(0.005)	(0.006)		(0.006)	(0.006)
cognitive skills squared		0.066***	0.066***		0.057***	0.058***
		(0.004)	(0.004)		(0.004)	(0.004)
background variables			yes			yes
constant	yes	yes	yes	yes	yes	yes
Observations	5,196	5,196	5,196	5,196	5,196	5,196
R-squared	0.082	0.424	0.457	0.082	0.410	0.450

Note: Linear probability model of the probability to receive an advice for an academic track (7-9: havo-vwo). The additional background variables include demographic variables, family background, school characteristics and regional charactristics. The standard errors are clustered at school level, significant at *** p<0.01, ** p<0.05, * p<0.1

Table B.2: Decomposition of gap in of biniary school advice

VARIABLES	(1) initial advice	(2) initial advice	(3) final advice	(4) final advice
mean by parental edu:				
high educated	0.607	0.607	0.635	0.635
	(0.015)	(0.015)	(0.015)	(0.015)
not high educated	0.316	0.316	0.341	0.341
	(0.012)	(0.012)	(0.012)	(0.012)
difference	0.291	0.291	0.294	0.294
	(0.016)	(0.015)	(0.016)	(0.015)
explained	0.165***	0.209***	0.164***	0.207***
	(0.010)	(0.013)	(0.010)	(0.013)
	56.7%	71.8%	55.8%	70.4%
unexplained	0.126***	0.083***	0.130***	0.087***
	(0.015)	(0.014)	(0.015)	(0.014)
	43.3%	28.5%	44.2%	29.6%
explained:				
non-cognitive skills	0.008***	0.007***	0.008***	0.007***
	(0.002)	(0.002)	(0.002)	(0.002)
cognitive skills	0.163***	0.157***	0.160***	0.156***
	(0.010)	(0.010)	(0.010)	(0.010)
cognitive skills squared	-0.005*	-0.005*	-0.004*	-0.004*
	(0.003)	(0.003)	(0.002)	(0.002)
background		0.038***		0.039***
		(0.007)		(0.007)
school		0.011**		0.011**
		(0.005)		(0.005)
regional		0.001		-0.001
		(0.003)		(0.003)
Observations	5,196	5,196	5,196	5,196

Note: Oaxaca-Blinder decomposition of the probability to receive an advice for an academic track (7-9: havo-vwo). The additional background variables include demographic variables, family background, school characteristics and regional charactristics. The standard errors are clustered at school level, significant at *** p<0.01, ** p<0.05, * p<0.1

Table B.3: Determinants of school advice, including all skills items

VARIABLES	(1) initial advice	(2) initial advice	(3) final advice	(4) final advice
parents not tertiary educated	-0.777***	-0.734***	-0.749***	-0.710***
	(0.068)	(0.061)	(0.068)	(0.060)
non-cognitive skills	0.213***		0.209***	
	(0.029)		(0.030)	
cognitive skills	1.616***		1.591***	
	(0.028)		(0.029)	
performance (rel)		-0.113***		-0.124***
		(0.042)		(0.042)
behavior		-0.069		-0.064
		(0.048)		(0.047)
work attitude		0.301***		0.307***
		(0.039)		(0.039)
popularity		0.034		0.031
		(0.037)		(0.036)
self-confidence		0.071		0.094**
		(0.044)		(0.043)
motivation		0.078*		0.077*
		(0.043)		(0.043)
wellbeing with teacher		-0.073*		-0.080*
		(0.044)		(0.043)
wellbeing with peers		-0.069*		-0.073*
		(0.040)		(0.040)
independence		0.138***		0.142***
		(0.039)		(0.039)
no conflict		0.001		-0.005
		(0.056)		(0.054)
closeness		-0.009		-0.021
		(0.054)		(0.053)
figure composition		0.029***		0.027***
		(0.010)		(0.010)
exclusions		0.032***		0.034***
		(0.010)		(0.011)
number series		0.065***		0.068***
		(0.010)		(0.010)
categories		0.009		0.009
		(0.009)		(0.008)
analogies		0.033***		0.032***
		(0.007)		(0.008)
reading comprehension		0.013***		0.013***
		(0.001)		(0.001)
technical reading		0.014***		0.014***
		(0.001)		(0.001)
vocabulary		0.031***		0.031***
		(0.001)		(0.001)
math		0.011***		0.010***
		(0.001)		(0.001)
Constant	6.108***	-0.587	6.226***	-0.399
	(0.063)	(0.357)	(0.061)	(0.355)
Observations	5,196	5,196	5,196	5,196
R-squared	0.541	0.631	0.531	0.620

Note: The additional background variables include demographic variables, family background, school characteristics and regional charactristics. The standard errors are clustered at school level, significant at *** p<0.01, ** p<0.05, * p<0.1

Table B.4: Decomposition of gap in school advice (all skills items)

VARIABLES	(1) initial advice	(4) initial advice	(7) initial advice	(1) final advice	(4) final advice	(7) final advice
mean by parental edu:						
high educated	6.686	6.686	6.686	6.795	6.795	6.795
not high educated	4.98	4.98	4.98	5.132	5.132	5.132
difference	1.706	1.706	1.706	1.663	1.663	1.663
explained	0.972***	1.233***	1.233***	0.953***	1.200***	1.200***
	(0.062)	(0.071)	(0.071)	(0.061)	(0.070)	(0.070)
	57.0%	72.3%	72.3%	57.3%	72.2%	72.2%
unexplained	0.734***	0.473***	0.473***	0.710***	0.464***	0.464***
	(0.060)	(0.051)	(0.051)	(0.060)	(0.052)	(0.052)
	43.0%	27.7%	27.7%	42.7%	27.9%	27.9%
explained:						
non-cognitive skills	0.066***	0.045***		0.066***	0.045***	
	(0.013)	(0.012)		(0.013)	(0.012)	
cognitive skills	0.906***	0.880***		0.887***	0.869***	
	(0.059)	(0.057)		(0.058)	(0.057)	
background		0.214***	0.214***		0.202***	0.202***
		(0.031)	(0.031)		(0.031)	(0.031)
school		0.102***	0.102***		0.097***	0.097***
		(0.031)	(0.031)		(0.031)	(0.031)
regional		-0.008	-0.008		-0.015	-0.015
		(0.018)	(0.018)		(0.018)	(0.018)
performance (rel)			-0.007*			-0.007*
			(0.004)			(0.004)
behavior			-0.006			-0.005
			(0.006)			(0.006)
work attitude			0.051***			0.049***
			(0.011)			(0.011)
popularity			-0.001			-0.002
			(0.004)			(0.004)
self-confidence			-0.001			-0.001
			(0.002)			(0.002)
motivation			-0.003			-0.002
			(0.003)			(0.003)
wellbeing with teacher			0.002			0.003
			(0.002)			(0.002)
wellbeing with peers			0.001			0.001
			(0.001)			(0.001)
independence			0.023***			0.024***
			(0.007)			(0.007)
no conflict			-0.011			-0.010
			(0.008)			(0.007)
closeness			-0.002			-0.002
			(0.003)			(0.003)
figure composition			0.025***			0.024***
			(0.008)			(0.008)
exclusions			0.020***			0.022***
			(0.007)			(0.008)
number series			0.040***			0.043***
			(0.008)			(0.009)
categories			0.010			0.011
			(0.009)			(0.009)
analogies			0.035***			0.032***
			(0.011)			(0.012)
reading comprehension			0.158***			0.153***
			(0.019)			(0.019)
technical reading			0.043***			0.041***
			(0.012)			(0.011)
vocabulary			0.380***			0.377***
			(0.032)			(0.032)
math			0.169***			0.165***
			(0.021)			(0.021)
Observations	5,196	5,196	5,196	5,196	5,196	5,196

Note: Oaxaca-Blinder decomposition. The additional background variables include demographic variables, family background, school characteristics and regional charactristics. The standard errors are clustered at school level, significant at *** p<0.01, ** p<0.05, * p<0.1

Table B.5: Determinants of initial school advice, by parental education low

VARIABLES	(1) initial advice	(2) initial advice	(3) initial advice	(4) initial advice
parental education unknown	0.328	0.132		
	(0.376)	(0.364)		
parental education middle	0.957***	0.346***		
	(0.100)	(0.080)		
parental education high	2.347***	1.052***		
	(0.112)	(0.095)		
migrant 1st gen	-0.091	0.487***		
	(0.233)	(0.174)		
migrant 2nd gen	-0.172*	0.455***		
	(0.102)	(0.081)		
parental education low (binary)			-0.487***	-0.258***
			(0.086)	(0.077)
migration background				0.508***
				(0.075)
non-cognitive skills		0.207***	0.215***	0.140***
		(0.029)	(0.030)	(0.029)
cognitive skills		1.692***	1.721***	1.609***
		(0.030)	(0.028)	(0.031)
cognitive skills squared		0.154***	0.168***	0.163***
		(0.018)	(0.018)	(0.018)
background				yes
Constant	4.371***	4.793***	5.547***	12.653***
	(0.094)	(0.082)	(0.060)	(0.721)
Observations	5,196	5,196	5,196	5,196
R-squared	0.134	0.554	0.533	0.598

Note: in clolumns 1 and 2 the base group for parental education is low educated and the base group for migration background is native. The additional background variables include demographic variables, family background, school characteristics and regional charactristics. The standard errors are clustered at school level, significant at *** p<0.01, ** p<0.05, * p<0.1

Table B.6: Decomposition of difference by parental education low in school advice

VARIABLES	(1) initial advice	(2) initial advice	(3) final advice	(4) final advice
mean by parental edu:				
not low educated	5.936	5.936	6.061	6.061
	(0.065)	(0.065)	(0.063)	(0.063)
low high educated	4.275	4.275	4.452	4.452
	(0.105)	(0.105)	(0.108)	(0.108)
difference	1.661	1.661	1.609	1.609
	(0.105)	(0.105)	(0.107)	(0.107)
explained	1.174***	1.402***	1.158***	1.351***
	(0.066)	(0.088)	(0.065)	(0.089)
	70.7%	84.4%	72.0%	84.0%
unexplained	0.487***	0.258***	0.451***	0.258***
	(0.087)	(0.079)	(0.088)	(0.077)
	29.3%	15.5%	28.0%	16.0%
explained:				
non-cognitive skills	0.045***	0.029***	0.044***	0.029***
	(0.011)	(0.008)	(0.011)	(0.008)
cognitive skills	1.181***	1.104***	1.157***	1.087***
	(0.070)	(0.067)	(0.069)	(0.066)
cognitive skills squared	-0.052***	-0.051***	-0.043***	-0.042***
	(0.012)	(0.012)	(0.010)	(0.010)
background		0.261***		0.233***
		(0.052)		(0.053)
school		0.159***		0.146***
		(0.054)		(0.055)
regional		-0.101***		-0.102***
		(0.032)		(0.031)
Observations	5,196	5,196	5,196	5,196

Note: Oaxaca-Blinder decomposition. The additional background variables include demographic variables, family background, school characteristics and regional charactristics. The standard errors are clustered at school level, significant at *** p<0.01, ** p<0.05, * p<0.1

Table B.7: Determinants of school advice using population weights

VARIABLES	(1) ES	(2) EMS	(3) all	(4) EM	(5) EMS	(6) all
parents not tertiary educated	-1.746***	-0.888***	-0.551***	-1.714***	-0.867***	-0.546***
	(0.089)	(0.072)	(0.060)	(0.089)	(0.072)	(0.059)
with migration background	-0.299***	0.318***	0.451***	-0.240**	0.371***	0.502***
	(0.109)	(0.083)	(0.072)	(0.110)	(0.084)	(0.072)
non-cognitive skills		0.206***	0.136***		0.204***	0.135***
		(0.034)	(0.032)		(0.036)	(0.033)
cognitive skills		1.653***	1.528***		1.630***	1.505***
		(0.032)	(0.035)		(0.032)	(0.035)
cognitive skills squared		0.140***	0.149***		0.119***	0.127***
		(0.021)	(0.020)		(0.021)	(0.020)
background variables			yes			yes
constant	6.845***	5.952***	14.338***	6.937***	6.075***	15.158***
	(0.073)	(0.081)	(0.919)	(0.070)	(0.078)	(0.903)
Observations	5,196	5,196	5,196	5,196	5,196	5,196
R-squared	0.131	0.554	0.614	0.127	0.547	0.611

Note: The additional background variables include demographic variables, family background, school characteristics and regional charactristics. The standard errors are clustered at school level, significant at *** p<0.01, ** p<0.05, * p<0.1

Table B.8: Decomposition of difference in initial school advice using poplation weights

VARIABLES	(1) by education		(2) by education	
mean by parental edu:				
high educated	6.796		6.796	
not high educated	5.024		5.024	
difference	1.772		1.772	
total explained	0.901***	50.8%	1.221***	68.9%
	(0.065)		(0.079)	
total unexplained	0.871***	49.2%	0.551***	31.1%
	(0.081)		(0.064)	
explained:				
non-cognitive skills	0.036***	2.0%	0.024***	1.4%
	(0.010)		(0.007)	
cognitive skills	0.872***	49.2%	0.816***	46.0%
	(0.063)		(0.060)	
cognitive skills squared	-0.007	-0.4%	-0.007	-0.4%
	(0.006)		(0.007)	
background			0.312***	17.6%
			(0.038)	
school			0.060**	3.4%
			(0.027)	
regional			0.017	1.0%
			(0.019)	
Observations	5,196		5,196	

Note: The additional background variables include demographic variables, family background, school characteristics and regional charactristics. The standard errors are clustered at school level, significant at *** p<0.01, ** p<0.05, * p<0.1