The effect of track placement on cognitive and non-cognitive skills [1]

Lex Borghans, Roxanne Korthals and Trudie Schils

Maastricht University

Abstract: Tracking in education is used to tailor education to the capabilities and the needs of each child. If every child is assigned to the track that fits his needs best, one would expect that children at the margin would be indifferent between the two tracks at stake. The aim of this paper is to investigate the effect of being in the higher track for students at the margin for a wide set of outcomes, including both cognitive and non-cognitive outcomes. For the analysis we use a longitudinal dataset on cognitive and non-cognitive skill development in both elementary and secondary education in a Dutch region. We apply a fuzzy regression discontinuity design using the discontinuity in a test score and a teacher recommendation in the assignment to tracks. Our main finding is that track placement influences IQ, the reading skills development and the self-perceived probability to obtain the degree for the marginal student but has no effect on personality traits, other non-cognitive skills and mathematics. Track mobility does not counteract the initial track placement.

1. Introduction

Tracking in education is used to tailor education to the capabilities and the needs of each child. If every child is assigned to the program that fits his needs best, one would expect that children at the margin would be indifferent between the two tracks at stake. In practice however, parents and children tend to put in a lot of effort in getting into higher tracks. This suggests that, at least from their perspective, the high track is more attractive than the middle track for a larger group of students.

The aim of this paper is to investigate the effect of being in the high track for students at the margin for a wide set of outcomes, including both cognitive and non-cognitive outcomes. We apply a fuzzy regression discontinuity design (RDD; Imbens and Lemieux, 2007) using a discontinuity in a test score and the teacher recommendation in the assignment to tracks. Our main finding is that track placement influences IQ, the reading skills development and the self-perceived probability to obtain the degree for the marginal student but has no effect on personality traits, other non-cognitive skills and mathematics. Track mobility does not counteract the initial track placement.

[1] We would like to thank Bas ter Weel, participants of the Economics of Education group in Maastricht University, of the International Workshop on Applied Economics of Education 2014, of the CEPA PhD workshop (Stanford University) and of the AMCIS conference on Educational Systems in 2014 for useful comments.

For the analysis we use a longitudinal dataset on cognitive and non-cognitive skill development in both primary and secondary education in a Dutch region. In the Netherlands students are placed into tracks between 6th (elementary school) and 7th grade (secondary school). The used dataset contains the two main sources of information Dutch secondary schools receive from the elementary school to decide on track placement: the score of a uniform elementary school exit test and the elementary school teacher recommendation. We exploit these two signals to look at the marginal student. However, secondary schools differ somewhat in the assignment procedures they adhere to, which does not allow for a sharp RDD. Each school is free in its student acceptance policies, although all are required by law to use the two sources of information received from the elementary schools. To check for bias due to remaining endogeneity in the tracking decision, for a number of outcomes variables we additionally use available panel information. For several outcome variables we have similar measures in both before and after track placement.

This paper contributes to the literature on tracking, but is also closely related to issues on ability grouping and selective schools.[2] The literature on the effects of tracking, streaming, and ability grouping is very extensive and can be divided into papers which look at the effects of a substantial increase in the number of students entering the higher track or those looking at the marginal student who moves track.[3] The papers which look at a substantial inflow of lower ability students into the high track show, besides the tracking effect, also the effects of a changing composition of the high track since more lower ability peers are allowed into the higher track. Guyon, Maurin, and McNally (2012) and Van Elk *et al.* (2011) look at such an increased inflow of students into the high track in Northern Ireland and the Netherlands and find positive effects on outcomes of these students. Duflo, Dupas, and Kremer (2011) find, using an experiment in Kenya in which groups of students were assigned to a school with and without ability grouping, that ability grouping has positive overall effects on cognitive outcomes.

[2] Selective schools can be considered as the higher track, for instance when they prepare students for university entrance exams (ie. the so-called preparatory schools). These so-called preparatory schools are quite common in France, but exist also in the United States, United Kingdom and Canada.

[3] Some studies on tracking, like Hanushek and Woessmann (2006) and Ariga and Brunello (2007), focus on the overall effects of tracking which compare different tracking policies across countries. A number of other papers on tracking make use of different tracking policies within one country, often due to policy changes, to look at the effect of tracking. E.g. Pekkarinen (2008), Pekkarinen *et al.* (2013), and Hall (2012) who all find little effect. For ability grouping, Betts and Shkolnik (2000) find that only the grouped classes with average ability suffer from grouping, while there is no effect for the lower ability groups and a small positive effect for the high ability grouped classes. Figlio and Page (2002) find no negative effect of ability grouping for low-ability students and find some evidence they might even benefit from ability grouping.

This paper does not look at the effects of a substantial increase in the number of students going to the high track, but focusses on the marginal student who does or does not go to this high track. Consequently, this study and related studies are able to isolate the treatment effect of being in the high track on the individual student since the composition of the higher track hardly changes when the marginal student enters the high track. An example of a similar paper is Borghans *et al.* (2011). They show that the threshold in the Netherlands for the highest track is too high: Students below the threshold would benefit from being in the high track both in test scores and in later earnings. Dustmann *et al.* (2014) use month of birth as an instrument for track placement and show, using a reduced form, that month of birth has no effects on labor market outcomes. Pop-Echeles and Urquoia (2011) and Jackson (2010) use formal assignment rules in Romania and Trinidad and Tobago to instrument attendance of better achieving, or more selective, schools.[4] Both find that pupils in better schools have higher test scores at the end of secondary school. Jackson (2010) also finds that students in better schools pass more exams and more often earn a certificate that gives access to university, while Pop-Echeles and Urquiola (2011) also look at behavior aspects and find that better teachers sort into better schools, parents at those schools are more involved, children do more homework, and child's self-perception is more positive.

There is a growing literature which analyzes the relation between non-cognitive skills, for instance the big 5 personality traits or motivation, and student performance (e.g. Heckman and Rubinstein, 2001; Heller *et al.*, 2012). These non-cognitive skills are also shown to influence later outcomes (e.g. Heckman *et al.*, 2012; Heckman and Rubinstein 2001). However, notwithstanding this growing awareness of the importance of non-cognitive skills, little is known about the effects of education on non-cognitive skills. The contribution of this paper is that we look at the marginal student who is just able to go to the high track and we look at a wide set of both cognitive and non-cognitive outcomes, while we also use information on the same outcome variables before tracking has taken place.

The structure of this paper is as follows: Section 4.2 will elaborate on the dataset and the graphical analysis. The model and results are provided in Section 4.3. Section 4.4 concludes.

2. Data and graphical analysis

[4] See Hoekstra (2009) for similar analyses for entry into selective colleges.

The data used in this paper are the result of a cooperative project with schools, schools boards and municipalities in which almost all elementary and secondary schools in Zuid-Limburg, a region in the South of the Netherlands, participate. The data comprise the cohort of students that were in the 6th grade in 2009 (last grade of elementary school) and in the 9th grade in 2012 (third grade of secondary school). Students enter the tracked system in the 7th grade which comprises three main tracks, with some further subdivisions in mainly the lowest track.[5] In 2011, a little more than fifty percent of the students aged 15 attended the lowest track; another 20 percent the middle track and twenty-five percent of students was in the highest track (CBS, 2012, Figure 1.2.4). In this paper, we focus on the two upper tracks in which a total of 45 percent of students were enrolled. For the students in the sample, the data include extensive information, including non-cognitive skills, reading, math, and IQ test scores in both 6th and 9th grade.[6] The data also include the information on the elementary school exit test and the elementary school teacher recommendation which is necessary for our identification strategy. Finally, information on the socio-economic background of the student's parents and information about the school is available.

The dataset contains 9,124 students in 9th grade of secondary school, and for 5910 we also know in which track they were in 7th grade (the first grade of secondary school). We focus here on the top two tracks which gives us 1,067 in the high track and 2,151 in the middle track.[7] Of these 3,218 students for 42 students we miss both their elementary school exit test score and their elementary school track recommendation, leaving us with 2,117 in the middle track and 1,059 in the highest track. We use the full sample, and do not restrict our sample to those within a small bandwidth around the cutoff, to obtain more precision (Lee and Lemieux, 2010).

[5] The three tracks are VMBO, HAVO, and VWO. VMBO is preparatory middle-level vocational education which lasts 4 years, and consists of the sub-tracks pure practical education (pro), VMBO-basic profession-oriented, VMBO-middle management-oriented, VMBO-mixed, and VMBO theoretical. HAVO is higher general continued education and lasts 5 years. VWO is preparatory scholarly education and lasts 6 years. VWO is split into the sub-tracks athenaeum and gymnasium which are essentially the same, except that gymnasium students also have the courses Latin and/or Greek. Secondary schools with only students of a single track and schools with multiple tracks exist alongside each other, although the tracks could be separated across different school buildings. This is especially the case for the bottom track. In the first year of secondary school, or sometimes in the first two years, so-called bridge classes exist in which students of multiple tracks are grouped together, but these classes only rarely consist of more than two tracks.

[6] Not all children received the complete student questionnaire, resulting in a smaller sample for civic engagement and school well-being questions. Also, not all children took all tests or all test questions. Using IRT test scores are put on the same scale for all children who saw 13 or more test questions on each of the tests. We use the expected posterior estimates using a 2 parameter Bayesian Markov Chain Monte Carlo model.

[7] The remaining 2692 students are in the bottom track or among the 167 students who are in the upper two tracks but who repeated the 7th grade. Of this last group we have no elementary school data (and for 11 students we have also no elementary school exit score) and since these students entered the school the year before, the threshold which they faced was different from the threshold of the other students. For these reasons we dropped them.

Table 1 shows the descriptive statistics on some key variables, separated for students in the highest and the middle track.[8] Students in the two tracks differ in some respects: compared to students in the middle track students in the highest track not only have higher IQ and higher reading and math test scores, they also have higher perseverance, social skills, are more open, have a higher self-perceived probability to obtain a secondary school degree, are less positive about their labor market chances, and have higher educated parents. To see whether these differences occur due to selection or due to being in the high track is the goal of this paper.

Acceptance and track placement of students in secondary school is guided by the Dutch government: Each elementary school is required to send to the preferred secondary school of the student the elementary school teacher recommendation for track placement, and a second independent and objective measure (Kingdom of the Netherlands, 1981). To obtain this independent and objective measure at the end of the last grade in elementary school almost all students take a centralized exit test (the so-called CITO test).[9] The elementary exit test score ranges from 500 to 550 and the guidelines for the highest track state that a score of 538 is needed to go to the highest track and a score of 533 to go to the second highest track (CITO Score, 2014). The mean test score in the highest track in our sample is 547 and for the middle track 540, with considerable variation.

[8] Appendix A provides the items on which the variables are based.
[9] It is not prescribed which independent and objective measure is needed and thus multiple elementary school exit tests are used. However, for eighty-five percent of schools this second objective measure is the CITO test score (CITO, 2014).

Table 1: Descriptive statistics of students in the 9th grade

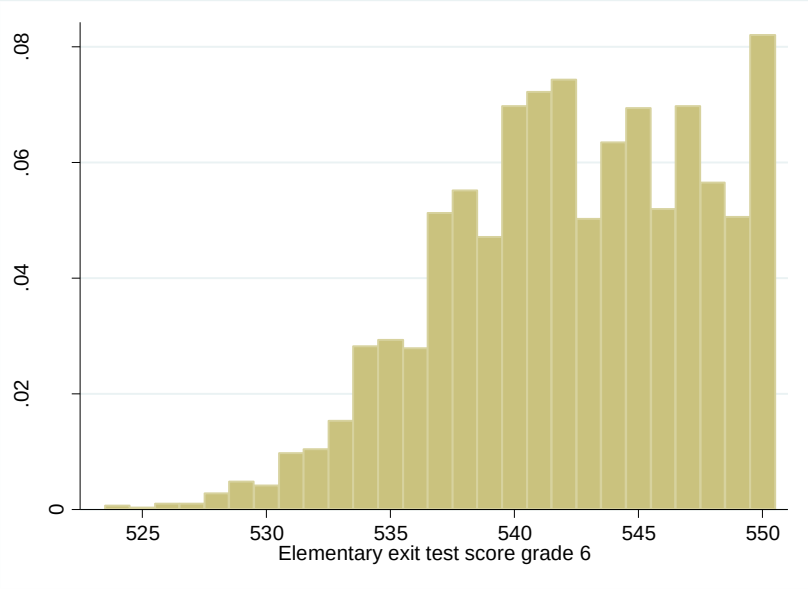
	Middle track (HAVO)					Highest track (VWO)						Dif in
	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max	Total obs	means
Age	1618	181.28	5.03	164.00	206.00	895	180.36	4.98	160.00	200.00	2513	0.92***
Gender	1585	0.54	0.50	0.00	1.00	888	0.50	0.50	0.00	1.00	2473	0.04*
Parental education	1447	2.52	1.03	1.00	4.00	811	2.89	1.01	1.00	4.00	2258	-0.37***
Elementary school exit test score	1618	540.40	4.91	524.00	550.00	895	546.01	3.72	527.00	550.00	2513	-5.61***
Recommendation	1618	15.84	1.83	4.00	19.00	895	17.81	1.08	3.00	19.00	2513	-1.97***
IQ	644	-0.23	0.98	-3.70	1.97	531	0.30	0.90	-3.70	1.97	1175	-0.53***
Math[a]	1197	0.26	0.79	-2.15	2.75	665	0.60	0.88	-2.01	3.35	1862	-0.34***
Reading[a]	1244	0.11	0.82	-3.03	2.09	683	0.54	0.97	-3.06	2.24	1927	-0.43***
Track grade 9	1618	5.14	0.78	1.00	6.00	895	5.86	0.37	4.00	6.00	2513	-0.72***
Openness	515	0.59	0.15	0.13	0.94	439	0.64	0.16	0.13	1.00	954	-0.05***
Conscientiousness	515	0.53	0.17	0.00	1.00	439	0.52	0.18	0.00	0.95	954	0.01
Extraversion	515	0.58	0.18	0.06	1.00	437	0.57	0.20	0.00	1.00	952	0.01
Agreeableness	515	0.74	0.16	0.00	1.00	436	0.75	0.17	0.00	1.00	951	-0.01
Neurotism	515	0.40	0.21	0.00	1.00	438	0.38	0.21	0.00	1.00	953	0.01
Perseverance	513	0.65	0.18	0.13	1.00	436	0.67	0.17	0.06	1.00	949	-0.02**
Competitive spirit	513	0.88	0.15	0.25	1.00	432	0.87	0.16	0.00	1.00	945	0.01
Social skills	499	0.62	0.09	0.00	0.90	433	0.63	0.08	0.25	0.90	932	-0.01***
Soc. sk: Social	498	0.60	0.09	0.00	0.90	433	0.62	0.09	0.24	0.90	931	-0.02***
Soc. sk.: Action	499	0.64	0.11	0.00	0.90	431	0.64	0.10	0.14	0.90	930	0.00
Self- p prob of obt degree	496	4.22	0.91	1.00	5.00	424	4.34	0.83	1.00	5.00	920	-0.12**
Prob of job 1	111	3.35	0.96	1.00	5.00	54	3.02	0.84	1.00	4.00	165	0.33
Prob of job 2	111	4.09	0.71	2.00	5.00	54	4.11	0.60	2.00	5.00	165	-0.02
Sch satisfaction	1367	6.49	1.36	1.00	10.00	723	6.61	1.38	1.00	10.00	2090	-0.11*
School motivation	1402	0.68	0.13	0.13	1.00	754	0.72	0.13	0.13	1.00	2156	-0.04***

Notes: How the outcome variables are defined is discussed in Appendix A. [a] Students made a math and a reading test, but not all students had the same questions. To ensure all students receive a test score on the same scale we used IRT to rescale the test scores. In italics, the variables for which there is a significant difference between children in the two tracks. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Figure 1 shows the density of the test score and the elementary school recommendation for students entering in the top two tracks. From Figure 1a it is clear that a ceiling effect occurs: The density of the test score is negatively skewed and many students get scores in the top range of the scale. The same can be seen in Figure 1b for the elementary school teacher recommendation, although to a lesser extent. Figure 2 shows that both test score and the elementary school recommendation clearly influence track placement, as the probability to enter in the highest track increases with both measures. Since there is no predefined cut off point, we apply a fuzzy RDD which assumes that, although the probability to enter in the highest track does not jump to 1 after the cut off, the probability increases for larger values of the forcing variable (Imbens and Lemieux, 2007).[10] Using a fuzzy RDD we essentially instrument track placement by passing the threshold of the forcing variables: the elementary school exit test score and the elementary school teacher recommendation.

Figure 1: Density of forcing variables

1a: Elementary school exit test score



1b: Elementary school teacher recommendation

[10] Given our fuzzy RD and the lack of a (predetermined) cutoff we do not have to worry about students trying to manipulate their score to be above the cutoff. It is in all students' best interest to have the highest possible exit test score and teacher recommendation.

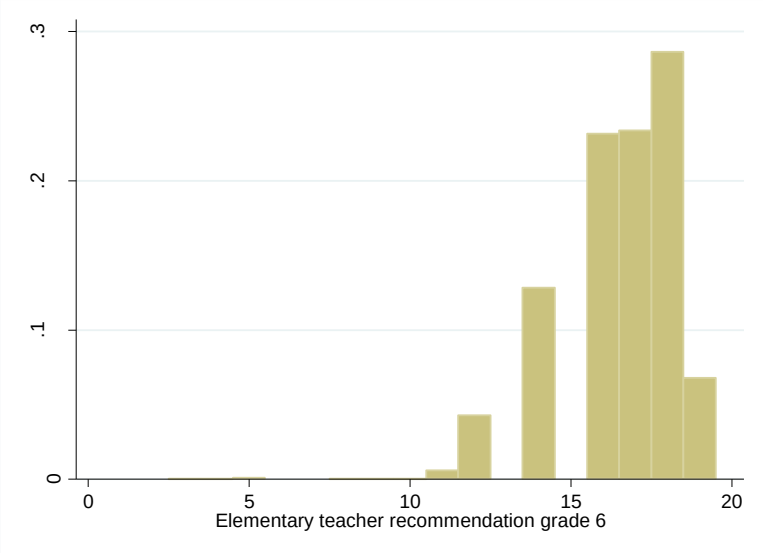
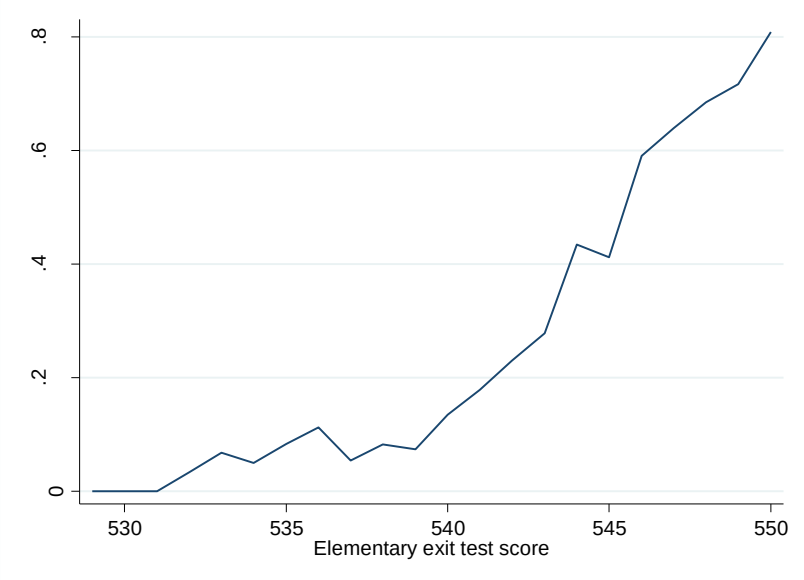
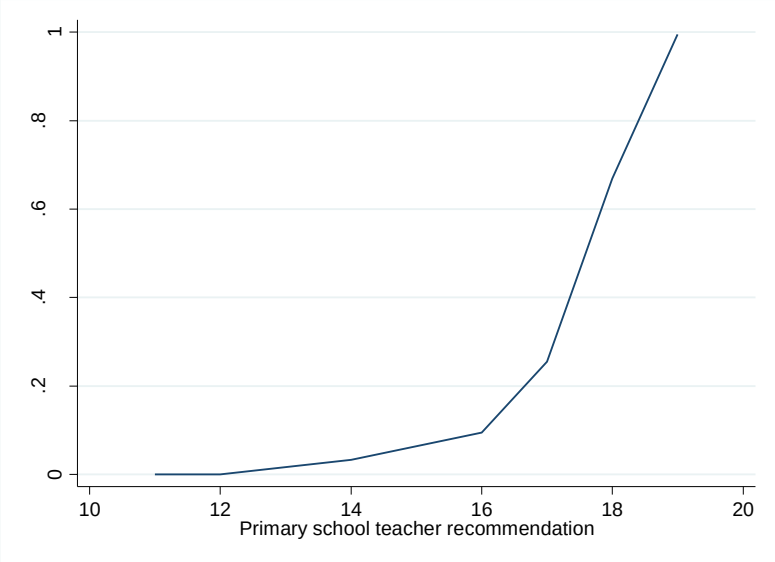


Figure 2: Fraction of being in the highest track over the forcing variables

1a: Elementary school exit test score



1b: Elementary school teacher recommendation

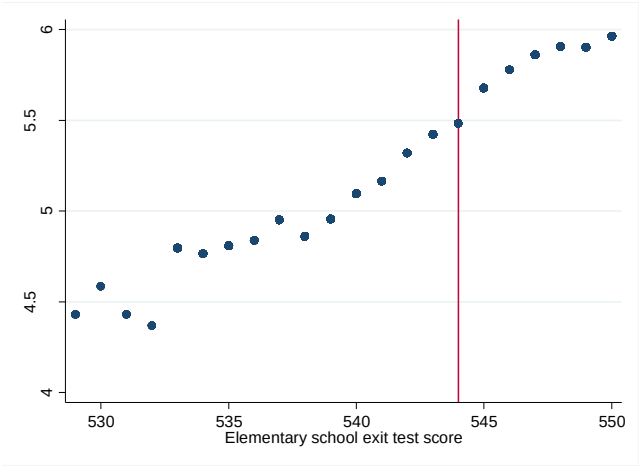


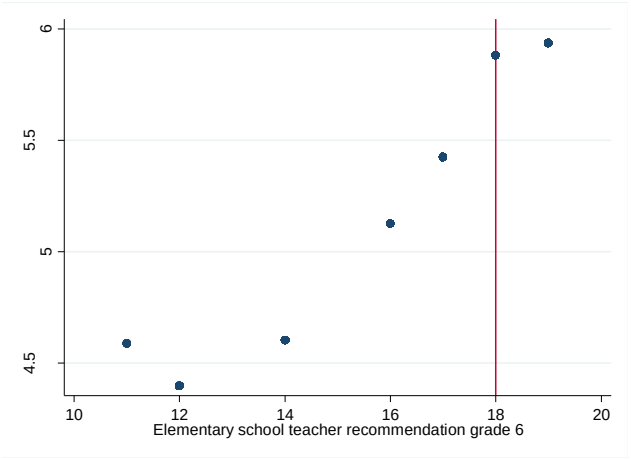
Notes: Only students with at least a reading or math score are included. Primary exit test scores which less than 10 students scored are not included.

Figure 3 shows for a number of outcome variables the average value over our two forcing variables: the exit test score and the elementary school teacher track recommendation. In these figures we see a clear semi-linear trend for the track in grade 9 (Figure 3a) and the reading test scores (Figure 3b) over the exit test score and the elementary school teacher recommendation. This trend is less clear for extraversion and social skills (Figure 3c and 3d). Since we use a fuzzy RDD, these figures are not very informative about the effects of being in the higher track for the marginal student. Using both the discontinuity in the test score and the elementary school recommendation, Section 4.3 shows that track placement has an effect on a number of outcome variables.

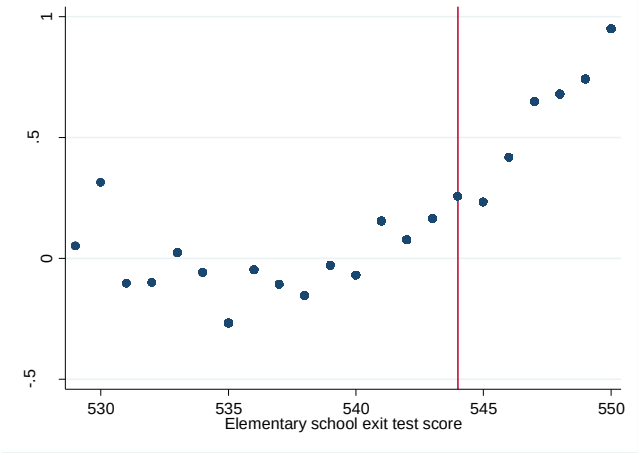
Figure 3: Average outcome values over the forcing variables

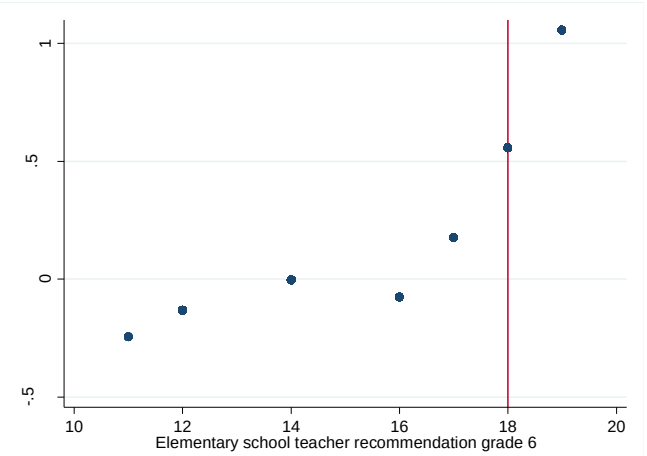
3a: Track in grade 9



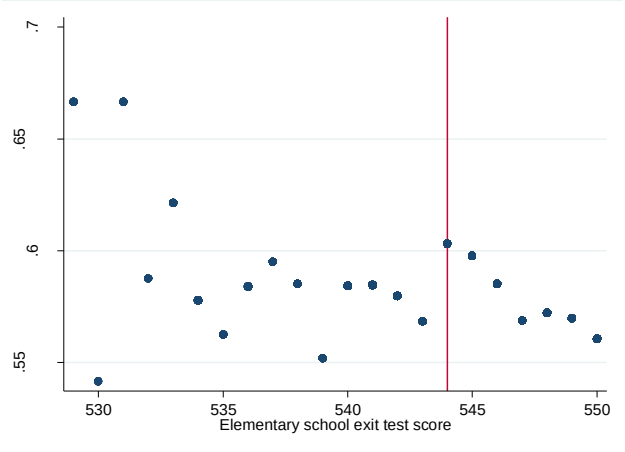


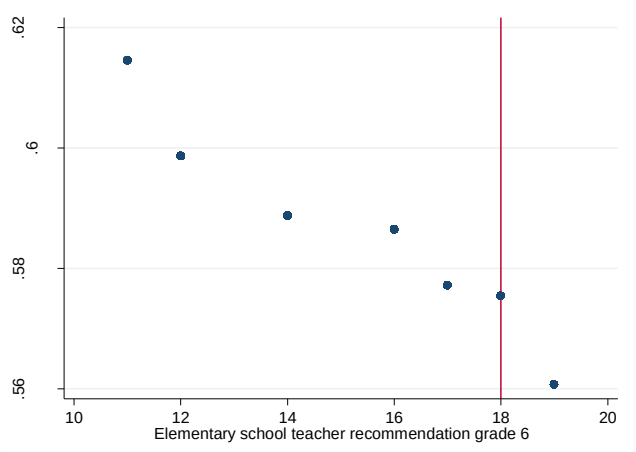
3b: Reading score in grade 9



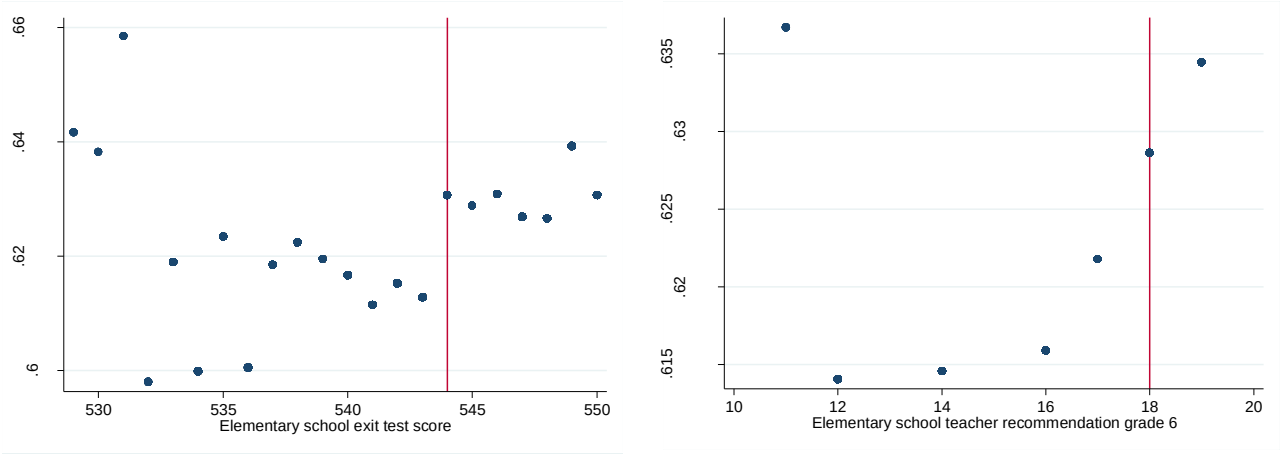


3c: Extraversion in grade 9





3d: Social skills in grade 9



Notes: Vertical bar is the threshold. The figures are only for those elementary school exit test scores and elementary school teacher recommendations which more than 10 students received.

Figures similar to Figure 3 to illustrate that background characteristics of the students do no influence track placement, or in other words that the students are identical around the threshold, are equally uninformative. To tackle this we use our IV strategy to predict some background characteristics, and Table 2 shows that we fail to do so showing that students are identical around the threshold.

Table 2: Identical students around the threshold

Dep var:	Gender	Age	Parental edu	Work Father	Work mother	Traditional family	IQ
High Track	0.09	-2.62	0.65	0.11	-0.12	-0.21	-0.09
	(0.14)	(2.11)	(0.48)	(0.20)	(0.19)	(0.18)	(0.15)
Test score	-0.01**	0.088	-0.01	-0.00	0.01	0.01	0.07***
	(0.01)	(0.07)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Recommendation	0.02	-0.02	0.04	-0.04*	0.01	0.02	0.03
	(0.01)	(0.29)	(0.04)	(0.03)	(0.03)	(0.02)	(0.02)
Constant	5.86**	135.10***	8.60	2.97	-2.43	-3.88	-35.56***
	(2.80)	(39.34)	(10.23)	(3.98)	(4.12)	(3.98)	(5.45)
# of students	1,175	1,175	923	911	817	970	1,175
# of schools	17	17	17	17	17	17	17
R^2	-	-	0.01	-	0.00	-	0.13

F 1st stage	15.25	15.25	9.72	8.44	6.42	8.56	15.25

Notes: The table presents coefficients (robust standard errors in parentheses) from IV models using the IQ sample. The first stage results are not shown for these samples, but is shown in Table 4 for the IQ sample. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

For some outcome variables (all measured in 9th grade) we have similar information available from 6th grade, the year before the students are being tracked. The panel dimension of our data is illustrated in Table 3, which provides descriptive statistics of the same variables as in Table 1, but now measured in elementary school in 6th grade. Again these statistics are separated by track, but at this age the students were still grouped together and the division into middle track versus highest track here is therefore merely an illustrative division. In 6th grade, the students who later entered the highest track had higher IQ, were more open, more agreeable, and had more social skills. However, they were also less extraverted and felt less at home at school. Between 6th and 9th grade we see that personality and school related measures change for students. This personality change over time is found more often (e.g. Roberts *et. al*, 2006), but it is yet unclear whether it is due to age differences in personality or due to changing environments over time, for instance entering in a new school or school type. In the last column of Table 3 the significant difference for students in the two tracks between the difference in values between 6th and 9th grade are shown. For instance, openness for students in the middle track was 0.02 points higher in 9th grade compared to 6th grade, while for students in the highest track this difference is -0.01. Students in the highest track became less open between 6th and 9th grade compared to students in the middle track, and this difference is significant at a 2%-significance level.

Table 3: Descriptive statistics of students in the 6th grade

	Middle track (HAVO)					Highest track (VWO)							
	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max	Total obs	Diff in means	P-value diff grade 6-9 by tracks
IQ	644	-0.18	0.98	-4.20	2.01	531	0.24	0.98	-4.52	2.01	1175	-0.43***	0.11
Openness	515	0.57	0.16	0.06	1.00	439	0.65	0.16	0.19	1.00	954	-0.08***	0.02
Conscientiousness	513	0.87	0.14	0.38	1.00	432	0.88	0.13	0.38	1.00	945	-0.01	0.06
Extraversion	515	0.52	0.16	0.06	1.00	439	0.54	0.17	0.05	1.00	954	-0.02	0.01
Agreeableness	515	0.48	0.13	0.00	0.81	437	0.45	0.13	0.06	0.94	952	0.02***	0.34
Neurotism	515	0.64	0.10	0.19	0.94	436	0.65	0.09	0.44	1.00	951	-0.01**	0.70
Perserverance	515	0.40	0.19	0.00	1.00	438	0.38	0.20	0.00	1.00	953	0.02	0.61
Competitve spirit	513	0.52	0.10	0.13	0.88	436	0.52	0.09	0.25	0.75	949	0.00	0.11
Social skills	499	0.49	0.10	0.21	0.75	433	0.52	0.09	0.27	0.75	932	-0.03***	0.09
Soc. sk.: Social	498	0.50	0.11	0.07	0.75	433	0.54	0.10	0.18	0.75	931	-0.04***	0.01
Soc. sk.: Action	499	0.49	0.12	0.15	0.75	431	0.49	0.12	0.15	0.75	930	0.00	0.90
Prob of obt degree	496	3.10	0.55	1.00	4.00	424	3.11	0.52	1.00	4.00	920	-0.01	0.08

Notes: In italics the variables for which there is a significant difference between children in the two tracks. In bold the variables for which there is a significant difference between the change between grade 6 and grade 9. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The last column shows the p-values of the difference between grade 9 and grade 6 for students in the middle and those in the highest track.

3. Analyses

Our model combines the advantages of a fuzzy RD design with a panel dimension. Using the cut off observed in the data for both the 6th grade test score and the elementary school teacher recommendation, we apply a fuzzy RD design in which we instrument track placement in 7th grade by passing the threshold for the elementary school teacher recommendation, or advice, and test score to study a number of outcomes using equation (1) and (2a). However, unlike Pop-Echeles and Urquoia (2011) and Jackson (2010), who use formal assignment rules to instrument selective school attendance, in the Netherlands no centralized cut off point is set. Schools are obliged to base their track placement decision on the elementary school exit test and the elementary school teacher recommendation, but each school is free to set its own cut off point with regards to its supply of students. We therefore instrument track placement in 7th grade by the two signals secondary schools receive to decide on track placement (Imbens and Lemieux, 2007). Some remaining endogeneity may still exist, for instance when schools deviate from the placement guideline for the test score and the elementary school teacher recommendation with reason. Therefore we also use the panel dimension of this data to limit our measurement error and remove any remaining selection: By controlling for the grade 6 outcome variable we only make use of the change in the outcomes variable due track placement (equation 1 and 2b).

Since there is no official elementary school exit test thresholds for which above the student automatically goes to the higher track, we use the test score for which we find the strongest link between track placement in 7th grade and an indicator function of having a test score above the cut off.[11] The analysis reveals that 544 is the unofficial cut off as seen in the data, and we subsequently use this cut off as if it was the cut off used by schools. For the elementary school teacher recommendation we use as cut off the recommendation that states that the child should go to the highest track.[12]

We estimate the following model:

$$HighTrack_i = \gamma + \delta_1 I(TEST \geq 544)_i + \delta_2 I(RECOM \geq 18)_i + \boldsymbol{X_i}\theta + \epsilon_i \qquad (1)$$

[11] The cut-off with the strongest link between track placement and the indicator function is the cut-off for which the F statistic reveals the strongest link.

[12] There are actually two categories that related to a elementary school teacher recommendation of the highest track. Recommendation 18 refers to the VWO-athenaeum, and recommendation 19 refers to the VWO-gymnasium, or bilingual education. A elementary school teacher recommendation for a bridge class of HAVO and VWO (the two upper tracks) is categorized as a recommendation for the middle track.

$$Y_{i,t=9} = \alpha + \beta_1 \widehat{HighTrack}_i + \beta_2 TEST_i + \beta_2 RECOM_i + \varepsilon_i \quad (2a)$$

$$Y_{i,t=9} = \alpha + \beta_1 \widehat{HighTrack}_i + \beta_2 TEST_i + \beta_2 RECOM_i + \beta_3 Y_{i,t=6} + \varepsilon_i \quad (2b)$$

where $HighTrack_i$ is an indicator whether the student was placed into the high track in 7th grade and is estimated in equation (1) and the fitted values from Equation (1) ($\widehat{HighTrack}_i$) are used as an explanatory variable in Equation (2a) and (2b). $HighTrack_i$ is estimated using the two data thresholds: the test score and the elementary school teacher recommendation. The matrix $\boldsymbol{X_i}$ contains the included instruments, i.e. the two forcing variables when using Equation (1a) and also the 6th grade outcome variables when using Equation (1b). $Y_{i,t}$ is an outcome variable in grade t=6 or t=9 (for instance a reading test score or extraversion), and $TEST_i$, the individual test score, and $RECOM_i$, the elementary school teacher recommendation for the student, are our two running variables. In Appendix B, we supplement this main model with models where either the test score or the elementary school teacher recommendation is used as an instrument, or where an interaction term is included for students who have both a test score and an elementary school teacher recommendation above the threshold. These models show qualitatively similar results.

The first stage is depicted in the first column of Table 4 and shows that students with higher test scores and with higher elementary school teacher recommendations are more likely to be in the high track as is to be expected. Our two instruments (having a test score of greater or above 544 and an elementary school teacher recommendation greater or above 18) are both highly significant in predicting track placement in 7th grade. Judging by the amount of explained variation the test score has more predictive power than the elementary school teacher recommendation on its own, and together they have the most explanatory power with an F statistics of 15, well above the required F statistic of 10 as proposed by Staiger and Stock (1997) and later refined by Stock and Yogo (2005). Depending on the dependent variable in the second stage the sample will change, and subsequently also the corresponding F statistic of the first stage changes. For that reason, all tables will also include the F statistic of the excluded instruments. We also estimated models with either the test score (F statistic of 18) or the elementary school teacher recommendation (F statistic of 9) as instrument, or where an interaction term is included for students who have both a test score and an elementary school teacher recommendation above the threshold (F statistic of 10). See Appendix B for these results.

Table 4: First stage results following equation (1a).

	High track	High track
I(Test score ≥ 544)	0.13***	0.13***
	(0.05)	(0.05)
I(Recommendation ≥ 18)	0.34**	0.34**
	(0.15)	(0.15)
Test score	0.01**	0.01*
	(0.00)	(0.00)
Recommendation	0.06**	0.06**
	(0.02)	(0.02)
Grade 6 IQ		0.00
		(0.01)
Constant	-4.45**	-4.33**
	(1.79)	(1.76)
# of students	1,175	1,175
# of schools	17	17
R^2	0.44	0.44

Notes: The table presents first stage coefficients (robust standard errors in parentheses) from IV models using the IQ sample. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 5 and 6 shows the results for the cognitive outcomes in 9th grade using OLS and the RD approach. The OLS results suggest that track placement affects the 9th grade track position of students, while for IQ the coefficient is significant only at the 10% level. Using RDD, we also find no effect of track placement for math, but we do find that track placements affects reading, and similarly to OLS, although a much stronger effect, the track in 9th grade. The significance level for the effect of IQ is just above 5 percent (p=0.053). Table 7 and 8 presents results for the effects of track placement on non-cognitive skills. Using OLS we find that track placement does not lead to any differences in any of our non-cognitive outcomes for the marginal students, except for school motivation. This suggests that the placement procedures of schools are able to correctly place students into tracks. If we remove the endogeneity, we find an effect for the self-perceived probability of obtaining the degree.[13] The reason that the OLS and IV estimates differ is because the IV estimates capture the local average treatment effect for those students affected by our instrument while the OLS

[13] The non-cognitive skills for which no effect with OLS or RD is found are extraversion, conscientiousness, openness, neurotism, agreeableness, competitive spirit, perseverance, social skills, school motivation, and school satisfaction.

estimates depicts average differences, both controlling for the forcing variables (Lee and Lemieux, 2010). The downward bias of the OLS estimates could be explained by the role the forcing variables take: In the IV models we see that removing the endogeneity of being in the higher track shifts part of the effect of the forcing variables (as seen in the OLS models) to the dummy for being in the higher track. Part of the being in the high track effect is in the OLS model wrongly attributed to the fact that on average these students also perform better in general. We correct for this using IV.

Table 5: OLS: The effects of being in the higher track on cognitive outcomes in grade 9

Dep var:	IQ	Math	Reading	Track grade 9
High track	0.20*	0.03	0.10	0.25**
	(0.10)	(0.09)	(0.09)	(0.09)
Test score	0.04***	0.03***	0.04***	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)
Recommendation	0.05*	0.07***	0.04**	0.14***
	(0.02)	(0.02)	(0.02)	(0.02)
Constant	-21.92***	-17.99***	-24.23***	-16.26***
	(5.04)	(2.50)	(2.94)	(3.39)
# of students	1,175	1,862	1,927	2,866
# of schools	17	22	22	22
R^2	0.13	0.11	0.13	0.43

Notes: The table presents coefficients (robust standard errors in parentheses) from OLS models depicting the relation between cognitive outcomes and being in the higher track. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 6: IV: The effects of being in the higher track on cognitive outcomes in grade 9

Dep var:	IQ	Math	Reading	Track grade 9
High track	0.44*	0.31	0.86**	0.48***
	(0.23)	(0.23)	(0.35)	(0.16)
Test score	0.03***	0.02***	0.02**	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)
Recommendation	0.02	0.05*	-0.02	0.13***
	(0.03)	(0.03)	(0.03)	(0.03)
Constant	-17.79***	-13.78***	-12.05**	-12.59***
	(5.48)	(2.83)	(5.18)	(3.35)
# of students	1,175	1,862	1,927	2,866
# of schools	17	22	22	22
R^2	0.11	0.10	0.01	0.42
F excluded instruments	15.25	11.80	17.11	19.64

Notes: The table presents coefficients (robust standard errors in parentheses) from IV models depicting the effect of being in the higher track on cognitive outcomes. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The first stage results are presented for the IQ sample in Table 4.

Table 7: OLS: The effects of being in the higher track on non-cognitive outcomes in grade 9

Dep var:	Self-perceived probability for completing degree	Extraversion	Conscientiousness	School motivation
High track	-0.04	0.01	-0.01	0.02**
	(0.07)	(0.01)	(0.01)	(0.01)
Test score	0.01	-0.00	-0.00	0.00***
	(0.01)	(0.00)	(0.00)	(0.00)
Recommendation	0.04	-0.00	0.01*	0.00
	(0.03)	(0.01)	(0.00)	(0.00)
Constant	-3.57	1.95**	1.65	-0.76*
	(4.28)	(0.89)	(1.02)	(0.41)
# of students	994	1,026	1,028	2285
# of schools	16	16	16	22
R^2	0.02	0.00	0.00	0.04

Notes: The table presents coefficients (robust standard errors in parentheses) from OLS models depicted the relation between non-cognitive outcomes and being in the higher track. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 8: IV: The effects of being in the higher track on non-cognitive outcomes in grade 9

Dep var:	Self-perceived probability for completing degree	Extraversion	Conscientiousness	School motivation
High track	0.60**	0.03	-0.02	0.03
	(0.25)	(0.03)	(0.03)	(0.03)
Test score	-0.01	-0.00	-0.00	0.00**
	(0.01)	(0.00)	(0.00)	(0.00)
Recommendation	-0.03	-0.00	0.01**	0.00
	(0.04)	(0.01)	(0.00)	(0.00)
Constant	7.46	2.35*	1.53	-0.53
	(5.32)	(1.23)	(1.16)	(0.49)
# of students	994	1,026	1,028	2285
# of schools	16	16	16	22
R^2	-	0.00	0.00	0.11
F excluded instruments	11.42	10.64	10.89	12.42

Notes: The table presents coefficients (robust standard errors in parentheses) from IV models depicting the effect of being in the higher track on non-cognitive outcomes. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The first stage results are presented for the IQ sample in Table 4.

For part of our independent variables the same or similar variables are available for 6th grade, so we can include these as controls. These results are presented in table 9 for IQ and the self-perceived probability of obtaining the degree. The other variables showed insignificant results. Unfortunately we do not have test scores for either math or reading in 6th grade. Like before, the OLS results suggest an effect on IQ but not on the self-perceived probability to obtaining the degree. Table 9 shows that for IQ and for the self-perceived probability of obtaining the degree there is an effect of track placement. For IQ this effect is quite large and about twice as large as the OLS result: Being placed in the higher track leads to an increase in the IQ score of half a standard deviation. So only because a child is placed in the higher track in 7th grade, this child has a higher IQ score in 9th grade and think more positively of its chances of obtaining a secondary school diploma.

Table 9: The effects of being in the higher track on outcomes in grade 9, controlled for characteristics in grade 6.

Dep var:	IQ		Self-perceived probability for completing degree	
Method:	OLS	IV	OLS	IV
High track	0.19**	0.47**	-0.04	0.57**
	(0.09)	(0.22)	(0.07)	(0.23)
Test score	0.02**	0.02*	0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Recommendation	0.04*	0.02	0.04	-0.02
	(0.02)	(0.02)	(0.03)	(0.04)
Grade 6 variable	0.25***	0.25***	0.14***	0.15***
	(0.03)	(0.03)	(0.04)	(0.04)
Constant	-13.46**	-8.94*	-3.80	6.55
	(4.68)	(4.93)	(4.22)	(5.20)
# of students	1,175	1,175	994	994
# of schools	0.18	17	0.03	16
R^2	17	0.17	16	-
F excl. instr.	-	15.17	-	10.90

Notes: The table presents coefficients (robust standard errors in parentheses) from OLS and panel IV models depicting the effect of being in the higher track on IQ and the self-

perceived probability for completing degree. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Heterogeneity in the results exists for gender and is shown in Table 10.[14] The effect of track placement on IQ seems to come entirely from the boys, since the girls' IQ scores are not affected by track placement. The reading score of both boys and girls is affected by being in the higher track, but for boys this effect is only significant at the 10 percent level.

Table 10: Heterogeneity effects of being in the higher track: boys vs girls

Dependent variable:	IQ	IQ	Reading	Reading
Sample:	Boys	Girls	Boys	Girls
High track	0.92***	0.14	0.80*	0.82***
	(0.28)	(0.29)	(0.46)	(0.31)
Test score	0.01	0.05***	0.02	0.03***
	(0.01)	(0.01)	(0.02)	(0.01)
Recommendation	0.02	0.02	-0.01	-0.02
	(0.04)	(0.03)	(0.04)	(0.03)
Constant	-7.65	-25.79***	-10.04	-15.72***
	(5.30)	(6.36)	(8.46)	(4.55)
# of students	552	623	892	1,004
# of schools	17	17	22	22
R^2	0.08	0.12	0.01	0.08
F excl. instr.	13.52	6.206	16.69	11.54

Notes: The table presents coefficients (robust standard errors in parentheses) from IV models separately for boys and girls. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

As said before, we have also used different model specifications as presented in Appendix B, where we used not both the elementary school exit test score and the elementary school teacher recommendation but either on. And we also included an interaction between the two indicator functions in the first stage or added an interaction between the two forcing variables in the second stage. The results are quite robust to these model specifications, although the F statistics varies over the models.

[14] No real heterogeneity exists for specific age groups. Young children seem to be more extraverted if they are placed into the highest track, but the F statistic is insufficient (5.8), perhaps due to the small sample size (N=498). The results of these split sample analyses are available on request.

Since we have not restricted our sample using a bandwidth around the threshold to go to the higher track, and to provide an extra check to ensure students characteristic difference around the cutoff do not influence the results, we also estimated our models using controls for gender and parental background (Imbens and Lemieux, 2007). These results, which are very similar to those presented in the main text, can be found in Appendix C.

4. Conclusion

In this paper we use a fuzzy RD design together with a panel dimension to separate the selection effects from the track placement effects for students in the two higher tracks in the Netherlands. We look at differences in student characteristics in 9^{th} grade using the two main sources of information Dutch secondary schools receive from the elementary school to decide on track placement for students in 7^{th} grade: the score of a uniform elementary school exit test and the elementary school teacher recommendation. We find that most of the significant differences between students in the upper two tracks are due to selection and not due to track placement: The track placement procedures of schools in the Netherlands are able to correctly select students into tracks for almost all non-cognitive outcomes, but less so for cognitive outcomes. We find that IQ and reading test scores are enhanced simply by being in the highest track, and that track placement also increases the likelihood of being in the higher track in 9^{th} grade. Math scores, on the other hand, are not affected by track placement. Furthermore, the only non-cognitive skill that is affected by track placement is the self-perceived probability of obtaining the secondary school degree: Students in the high track are more positive about their chances of graduating. Other non-cognitive skills as personality or social skills are not affected.

Although some parents strive for the highest track for their child due to the positive learning outcomes, other parents are hesitant to do so since they believe it might hamper their child's non cognitive development. We show that the non-cognitive development of the marginal student who goes to the high track is not affected by track placement, but that the child's cognitive development benefits from the higher track placement. Given that we find no negative effects for the marginal student to go to the high track, irrespective of the large number of insignificant results, it might be better if schools where more accommodating to accept the marginal student to the higher track.

The results presented in this paper do not imply that all students are better off in the high track. For the students who are not on the margin it might be beneficial for both cognitive and non-cognitive outcomes to remain in the low or middle track. Unfortunately, our set up does not allow us to investigate this.

References

Ariga, K., and Brunello, G (2007) Does Secondary School Tracking Affect Performance? Evidence from IALS. *IZA Discussion Paper* 2643.

Betts, J. and J. Shkolnik (2000) The effects of ability grouping on student achievement and resource allocation in secondary schools. *Economics of Education Review, 19*, pp. 1-15

Borghans, L., R. Diris, W. Smit and J. De Vries (2011) *The labor market effects of tracking: a regression discontinuity approach.* Thesis University Maastricht Ron Diris.

CBS (2012) Jaarboek onderwijs in cijfers 2012. Centraal Bureau voor de Statistiek: Den Haag/Heerlen.

CITO (2014) *Cito dossier Eindtoets 2014.* Arnhem: CITO.

CITO Score (2014) Schooladvies. http://www.Cito-score.nl/schooladvies.html. Last accessed 24-02-2014.

Clark, D. (2010) Selective Schools and Academic Achievement. The B.E. Journal of Economic Analysis & Policy Advances, 10(1), article 9.

Duflo, E. P. Dupas and M. Kremer (2011) Peer effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *American Economic Review*, 101, pp. 1739-1774.

Dustmann, C., P. Puhani and U. Schönberg (2014) The Long-Term Effects of Early Track Choice. *IZA Discussion Paper Series.* Working Paper 7897.

Figlio, D. and M. Page (2002) School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality? *Journal of Urban Economics*, 51, pp. 497-514.

Guyon, N., E. Maurin, S. McNally (2012) The effect of tracking students by ability into different schools. *Journal of Human Resources*, 47(3), pp. 684-721.

Hanushek, E. and L. Woessmann (2006) Does Educational Tracking Affect Performance and Inequality? Difference-in-Difference Evidence Across Countries. *Economic Journal, 116*, pp. C63-C76

Hall, C. (2012) The effects of reducing tracking in upper secondary school. *Journal of Human Resources*, 47(1), pp. 238-269.

Hoekstra, M. (2009) The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach. *The Review of Economics and Statistics,* 91(4), pp. 717–724.

Heckman, J.J., R. Pinto, and P.A. Savelyev (2012). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. NBER working paper no. 18581, Cambridge: NBER.

Heckman, J.J. and Y. Rubinstein (2001). "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." American Economic Review 91(2), pp. 145-149.

Heller, S., H. Pollack, R. Ander, and J. Ludwig (2013) Preventing Youth Violence and Dropout: A Randomized Field Experiment. *NBER working paper*, no. 19014, Cambridge: NBER.

Imbens, G. and T. Lemieux (2007) Regression Discontinuity Designs: A Guide to Practice. *Journal of Econometrics*, 142(2), pp. 615-635.

Jackson, K. (2010) Do better students benefit from attending better schools? Evidence from rule-based student assignments in Trinidad and Tobago. *The Economic Journal*, 120(December), pp. 1399–1429.

Kingdom of the Netherlands (1981) *Wet op het primair onderwijs.* Hoofdstuk I. Basisonderwijs. Artikel 42. Onderwijskundig rapport. Law book. The Netherlands.

Lee, D. and T. Lemieux (2010) Regression Discontinuity Designs in Economics, *Journal of Economic Literature*, 48(June), pp. 281-355.

Pekkarinen, T. (2008) Gender Differences in Educational Attainment: Evidence on the Role of Tracking from a Finnish Quasi-experiment. *Scandinavian Journal of Economics* 110(4), 807-825.

Pekkarinen, T., Uusitalo, R.,& Kerr, S. (2009) School Tracking and Development of Cognitive Skills. *IZA Discussion Paper* 4058.

Pop-Eleches, C. and M. Urquiola (2011) Going to a better school: Effects and behavioural responses. *NBER working paper series*, Working Paper 16886.

Roberts, B., K. Walton, and W. Viechtbauer (2006) Patterns of Mean-Level Change in Personality Traits Across the Life Course: A Meta-Analysis of Longitudinal Studies. *Psychological Bulletin*, 132(1), pp. 1–25.

Saavedra, J. (2009) The Learning and early labor market effects of college quality: A Regression Discontinuity Analysis. *Mimeo.*

Staiger, D. and J. Stock (1997) Instrumental Variables Regression with Weak Instruments. *Econometrica*, 65(3), pp. 557-586.

Stock, J. and M. Yogo (2005) Testing for Weak Instruments in Linear IV Regression. In: D. Andrews, J. Stock, T. Rothenberg (eds) Identification and inference for econometric models: Essays in honor of Thomas Rothenberg, pp. 80-108. New York: Cambridge University Press.

Van Elk, R., van der Steeg, M., & Webbink, D. (2011) Does the Timing of Tracking Affect Higher Education Completion? *Economics of Education Review* 30(5),1009-1021.

Appendix A- Outcome variables

Grade 7 and 9	
IQ 1	
	Number of IQ questions correct as percentage of total number of questions.
IQ 2	
	Number of IQ questions correct as percentage of total questions completed.
Openness	
	I use difficult words
	I have lots of ideas
	I learn things quickly
	I have little imagination
Competitive spirit	
	I would like to get high marks
	Later I want to be good at my job
Conscientiousness	
	I do my chores immediately
	I often leave my stuff hanging around
	I always stick to my appointments
	I sometimes forget I have to do something
	I am very precise in what I do
Extraversion	
	I talk a lot (negatively formulated in grade 7)
	I am quite among strangers
	I am the pacesetter at parties
	I like to be around lots of people
Agreeableness	
	I try to help people
	I am interested in others (negatively formulated in grade 7)
	I empathize with others
	I am a friendly person
Neurotism	
	I easy get upset
	I am often stressed
	My temper shifts often
	I regularly have a gloomy mood
Perseverance	
	I continue until it is done
	I stop easily if it gets too difficult
	If I start something, I finish it
	If something is harder than expected, I soon loose heart
Social skills	
Social skills (action)	Own appreciation for: drawing, painting or making music
	Own appreciation for: looking for something on the computer

	Own appreciation for: writing without mistakes
	Own appreciation for: mental arithmetic
	Own appreciation for: own ability to concentrate
	Own appreciation for: to choose nice clothes and look good
	Own appreciation for: own ability to discuss
Social skills (social)	Own appreciation for: own ability to comfort someone
	Own appreciation for: own ability to give your opinion
	Own appreciation for: own ability to win an argument
	Own appreciation for: own ability to get my way
	Own appreciation for: own ability to interact with other students
School wellbeing	
School wellbeing student	I like going to this school
	I hate this school
	I like it at this school
	I am bored at this school
School wellbeing teachers	The teachers like me
	If I want the help of a teacher, I also receive it
	The teacher think I am smart
	The teacher do their very best for me
School wellbeing classmates	I have here many friends
	Some students bully me
	The students here like me
Estimated probability of completing secondary school	
	What is your estimation that you will finished your current degree?/Do you think you will manage in your next school?
Only in grade 9	
Estimated probability of obtaining a job 1	
	How large do you think the probability is that you can find a job easily if you finish your current degree?
Estimated probability of obtaining a job 2	
	How large do you think the probability is that you can find a job easily if you finish your next degree?
Civic engagement: Democratic behavior	
	Test score
Civic engagement: International mind set	
	Test score
School motivation	
	I will drop out without finishing school
	As soon as possible, I'll stop learning
	I will learn a profession, but outside school
	I am very motivated to continue learning
	I am going to learn interesting things

	I am going to continue learning because I like it
	I am going to continue learning for a very long time
	As soon as I can get a job, I will drop out of school
School satisfaction	
	I feel secure at school
	The atmosphere at school is nice
	The teacher treat me with respect
	There are clear rules at school
	The variation is the teaching styles is large
	The teacher explain things well
	I learn a lot from the teachers
	The teachers take into account what I can and cannot do
	I am sufficiently challenged to do my best in school
	I am satisfied about my mentor
	I know who to turn to at school in case of problems
	If needed, I receive extra tutoring
	I am assisted in making important choices regarding my studies
	Teachers clearly tell me how my results are
	I am informed about things relevant for students
	The opinion of students counts at this school
Math test score 1	
	Test score
Math test score 2	
	Test score
Reading test score 1	
	Test score
Reading test score 2	
	Test score

Appendix B- Other model specifications

Supplementing the analyses done, in this Appendix we use different model specifications, where we not use not both the elementary school exit test score and the elementary school teacher recommendation but either one (model IV-1 and IV-2 in the tables below), or we add an interaction between the two in the second stage (model IV-5) and an interaction between the two indicator functions in the first stage (model IV-4). The results are very robust to model specification, although the F statistics varies over the models.

Table B1: The effect of being in the high track on IQ in grade 9

Model:	OLS	IV-1	IV-2	**MAIN IV**	IV-4	IV-5
High Track	0.195*	0.794	0.638***	0.444*	0.445*	0.192
	(0.101)	(0.485)	(0.205)	(0.229)	(0.229)	(0.388)
Test score	0.039***	0.020		0.032***	0.032***	-0.041
	(0.010)	(0.022)		(0.010)	(0.010)	(0.081)
Recommendation	0.045*		0.063*	0.023	0.023	-2.463
	(0.023)		(0.033)	(0.025)	(0.025)	(2.790)
Test score* Rec.						0.005
						(0.005)
Constant	-21.920***	-11.020	-1.312***	-17.790***	-17.770***	20.790
	(5.038)	(11.720)	(0.481)	(5.476)	(5.456)	(42.900)
# of students	1,175	1,175	1,175	1,175	1,175	1,175
# of schools	17	17	17	17	17	17
R^2	0.13	0.07	0.09	0.11	0.12	0.13
F excl. instr.	-	18.57	9.421	15.25	10.44	3.258

Notes: The table presents coefficients (robust standard errors in parentheses) from OLS and IV models using different specification depicting the effect of being in the higher track on IQ. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table B2: The effect of being in the high track on reading in grade 9

Model:	OLS	IV-1	IV-2	**MAIN IV**	IV-4	IV-5
High Track	0.098	0.950***	1.022***	0.859**	0.867**	0.314
	(0.085)	(0.266)	(0.357)	(0.353)	(0.355)	(0.519)
Test score	0.044***	0.014		0.023**	0.022**	-0.146
	(0.005)	(0.011)		(0.009)	(0.009)	(0.093)
Recommendation	0.041**		0.010	-0.016	-0.016	-5.787*
	(0.018)		(0.042)	(0.027)	(0.027)	(3.313)
Test score* Rec.						0.011*
						(0.006)

Constant	-24.230***	-7.790	-0.274	-12.050**	-11.920**	77.920
	(2.940)	(6.073)	(0.549)	(5.181)	(5.252)	(49.930)
# of students	1,927	1,927	1,927	1,927	1,927	1,927
# of schools	22	22	22	22	22	22
R^2	0.13			0.01	0.01	0.13
F excl. instr.	-	17.26	20.64	17.11	13.25	2.126

Notes: The table presents coefficients (robust standard errors in parentheses) from OLS and IV models using different specification depicting the effect of being in the higher track on reading scores. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table B3: The effect of being in the high track on the track in grade 9

Model:	OLS	IV-1	IV-2	**MAIN IV**	IV-4	IV-5
High Track	0.254**	0.747***	0.683***	0.482***	0.471***	0.505**
	(0.0900)	(0.185)	(0.179)	(0.162)	(0.164)	(0.243)
Test score	0.0353***	0.0491***		0.0290***	0.0293***	0.0342
	(0.00641)	(0.0103)		(0.00606)	(0.00609)	(0.0577)
Recommendation	0.144***		0.159***	0.126***	0.127***	0.312
	(0.0234)		(0.0337)	(0.0275)	(0.0276)	(1.998)
Test score * Rec.						-0.000349
						(0.00374)
Constant	-16.26***	-21.51***	2.498***	-12.59***	-12.76***	-15.41
	(3.390)	(5.539)	(0.488)	(3.346)	(3.366)	(30.83)
# of students	2,866	2,866	2,866	2,866	2,866	2,866
# of schools	0.432	0.333	0.368	0.417	0.418	0.414
R^2	22	22	22	22	22	22
F excl. instr.	-	20.43	23.34	19.64	16.30	2.054

Notes: The table presents coefficients (robust standard errors in parentheses) from OLS and IV models using different specification depicting the effect of being in the higher track on the track in grade 9. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table B4: The effect of being in the high track on the self-perceived probability of obtaining the degree in grade 9

Model:	OLS	IV-1	IV-2	**MAIN IV**	IV-4	IV-5
High Track	-0.0438	0.949**	0.496**	0.601**	0.603**	0.420
	(0.0744)	(0.411)	(0.228)	(0.248)	(0.248)	(0.451)
Test score	0.0134	-0.0314		-0.00558	-0.00564	-0.0586
	(0.00862)	(0.0237)		(0.01000)	(0.00998)	(0.111)
Recommendation	0.0367		-0.0208	-0.0256	-0.0258	-1.865

	(0.0326)		(0.0402)	(0.0393)	(0.0394)	(3.852)
Test score * Rec.						0.00345
						(0.00722)
Constant	-3.567	20.88*	4.400***	7.460	7.489	35.76
	(4.283)	(12.63)	(0.555)	(5.324)	(5.316)	(59.16)
# of students	994	994	994	994	994	994
# of schools	16	16	16	16	16	16
R^2	0.017					
F excl. instr.	-	17.35	8.203	11.42	7.769	4.071

Notes: The table presents coefficients (robust standard errors in parentheses) from OLS and IV models using different specification depicting the effect of being in the higher track on the self-perceived probability to obtain the degree. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Appendix C- Including controls for gender and parental background

This paper present OLS, RD and RD panel results without any controls (Tables 5 to 9). In this appendix controls are added for parental background and gender. Table C1 shows the first stage using the IQ sample, while Tables C2 and C3 show the second stage for the models looking at IQ, the reading score, the track in grade 9 and the self-perceived probability to obtain the degree, which are the four outcomes for which we find significant effects using IV. The results presented here are very similar to those in Tables 5 to 9 in the main text.

Table C1: First stage results controlling for gender and parental education

Controlled for:	Gender		Parental education	
Dependent variable:	High track	High track	High track	High track
I(Test score>=544)	0.14***	0.14***	0.14**	0.14**
	(0.04)	(0.04)	(0.05)	(0.05)
I(Recommendation>=18)	0.34**	0.33**	0.31*	0.31*
	(0.15)	(0.15)	(0.15)	(0.15)
Test score	0.01*	0.01*	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
Recommendation	0.06**	0.06**	0.05**	0.05**
	(0.02)	(0.02)	(0.02)	(0.02)
Grade 6 IQ		0.00		0.00
		(0.01)		(0.01)
Control	-0.02	-0.02	0.01	0.01
	(0.02)	(0.02)	(0.02)	(0.02)
Constant	-4.28**	-4.13**	-6.28***	-6.14***
	(1.80)	(1.77)	(1.59)	(1.80)
# of students	1,175	1,175	923	923
# of schools	17	17	17	17
R^2	0.44	0.44	0.43	0.43

Notes: The table presents first stage coefficients (robust standard errors in parentheses) from IV models using the IQ sample, controlling for either gender or parental education. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table C2: The effect of being in the high track on outcomes, controlling for parental education.

	IQ			Self-perceived probability to obtain the degree			Reading		Track grade 9	
	OLS	IV	Panel IV	OLS	IV	Panel IV	OLS	IV	OLS	IV
High track	0.20	0.44*	0.49*	-0.081	0.48**	0.45**	0.089	0.84*	0.24**	0.44**
	(0.12)	(0.24)	(0.26)	(0.08)	(0.23)	(0.22)	(0.10)	(0.44)	(0.09)	(0.19)
Test score	0.04***	0.03***	0.01	0.01	-0.00	-0.00	0.05***	0.03**	0.04***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Recommendation	0.05*	0.03	0.02	0.03	-0.02	-0.02	0.05*	-0.01	0.12***	0.11***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.04)	(0.02)	(0.03)
Grade 6 variable			0.28***			0.14***				
			(0.04)			(0.04)				
Parental education	0.04	0.03	0.03	0.08***	0.07***	0.07***	0.02	-0.01	0.07***	0.07***
	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.01)	(0.01)
Constant	-20.62***	16.36***	-6.17	-3.82	6.19	5.35	25.93***	-13.88**	16.85***	-13.62***
	(4.99)	(5.22)	(5.81)	(3.91)	(5.15)	(5.08)	(3.37)	(6.67)	(3.81)	(4.08)
# of students	923	923	923	947	947	947	1,580	1,580	2,186	2,186
# of schools	17	17	17	16	16	16	22	22	22	22
R^2	0.12	0.12	0.18	0.03			0.14	0.03	0.43	0.41
F excl instr.		9.74	9.71		12.54	11.81		10.34		12.17

Notes: The table presents coefficients (robust standard errors in parentheses) from OLS, IV and panel IV models depicting the effect of being in the higher track on outcomes, controlling for parental education. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table CF3: The effect of being in the high track on outcomes, controlling for gender.

	IQ	Self-perceived probability to	Reading	Track grade 9

				obtain the degree						
	OLS	IV	Panel IV	OLS	IV	Panel IV	OLS	IV	OLS	IV
High track	0.20*	0.44*	0.47**	-0.04	0.60**	0.57**	0.12	0.82**	0.27***	0.47***
	(0.10)	(0.23)	(0.22)	(0.07)	(0.24)	(0.23)	(0.09)	(0.37)	(0.09)	(0.15)
Test score	0.04***	0.03***	0.02*	0.01	-0.00	-0.00	0.04***	0.03***	0.04***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Recommendation	0.04*	0.02	0.02	0.04	-0.03	-0.03	0.04*	-0.01	0.14***	0.12***
	(0.02)	(0.03)	(0.02)	(0.03)	(0.04)	(0.04)	(0.02)	(0.03)	(0.02)	(0.03)
Grade 6 variable			0.25***			0.16***				
			(0.03)			(0.04)				
Gender	0.05	0.05	0.03	0.05	0.07	0.07	0.32***	0.34***	0.15***	0.16***
	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.03)	(0.03)
Constant	-22.12***	-18.09***	-9.11*	-3.85	7.05	6.06	-24.75***	-13.57**	-16.92***	-13.56***
	(4.95)	(5.41)	(4.81)	(4.22)	(5.17)	(5.05)	(3.17)	(5.39)	(3.40)	(3.26)
# of students	1,175	1,175	1,175	994	994	994	1,896	1,896	2,796	2,796
# of schools	17	17	17	16	16	16	22	22	22	22
R^2	0.13	0.12	0.17	0.02			0.16	0.07	0.44	0.43
F excl instr.		15.63	15.55		11.83	11.27		18.56		21.63

Notes: The table presents coefficients (robust standard errors in parentheses) from OLS, IV and panel IV models depicting the effect of being in the higher track on outcomes, controlling for gender. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.