De Economist (2024) 172:141–166
https://doi.org/10.1007/s10645-024-09436-4

ORIGINAL PAPER

Did Migrant Children Benefit from a Delay in the Dutch Primary School Exit Test?

Bart H. H. Golsteyn[1] · **Cécile Magnée**[2,3] · **Inge de Wolf**[4]

Accepted: 18 April 2024 / Published online: 25 June 2024

Abstract

This paper evaluates whether educational outcomes of first-generation migrant children improved relative to those of natives after a policy change which delayed an important primary school exit test by three months. Using Dutch register data and a difference-in-differences methodology, we show that the policy change increased the academic rank of migrants relative to natives upon first enrollment. The policy change, therefore, has had an important positive effect on the educational chances of migrant children. Our analyses suggest that the results are driven by higher relative exit test scores and higher relative teacher recommendations.

Keywords Primary school · Migrant students · Tracking · Teacher assessments · Education policy

JEL Classification I2 · I24 · J15

✉ Bart H. H. Golsteyn
b.golsteyn@maastrichtuniversity.nl

Cécile Magnée
c.a.j.magnee@ind.nl

Inge de Wolf
i.dewolf@maastrichtuniversity.nl

[1] Department of Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

[2] Sponsor: Research Centre for Education and the Labour Market, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

[3] Immigration and Naturalisation Service (IND), P.O. Box 16275, 2500 BG Den Haag, The Netherlands

[4] Research Centre for Education and the Labour Market, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

1 Introduction

Increased migration into developed countries has led to a substantial fraction of students with non-native backgrounds.[1] First-generation migrant children are among the most vulnerable children in the education system. Some of them are refugees, who may suffer trauma from experiencing war and the fleeing process, have a gap in their educational career, and lack a support network. Most countries attempt to integrate these students so that they thrive in education and society. Countries with tracked educational systems face specific challenges, because students with non-native backgrounds are more at risk of being tracked in lower educational levels than expected based on their competences (Burgess & Greaves, 2013).[2]

In the Netherlands, tracking takes place at age 12, when students reach the end of primary school, and is based on a primary school exit test and on the teacher's recommendation. In 2014, there has been a policy change regarding the timing of the primary school exit test. In the period before the policy change, students took the primary school exit test relatively early during the school year such that teachers could use this test score when they gave the secondary school track recommendation. After the policy change, from 2014–2015 onward, the exit test is taken three months later, and teachers give their secondary school track recommendation without knowing the child's primary school exit test score.

This paper evaluates whether this policy change regarding the timing of the primary school exit test changed the secondary school enrollment levels of first-generation migrant students relative to those of native students.[3] To analyze the effects of the policy change, we use a very rich register dataset that provides information on educational outcomes and background characteristics of the full population of the Netherlands. The data allow us to follow all children who attended primary school in the period 2011–2017 until they reach the third year in secondary school (i.e., in 2014–2021).[4] The data contain information on school enrollment levels, primary school exit test scores, and the primary school teachers' recommendations regarding secondary school levels. The data also contain characteristics of individuals, such as the year migrants entered the Netherlands. Additional benefits of having a dataset on the full population are that there is no bias due to self-reporting, sample-selection, or attrition.

[1] For instance, over the past years, the number of students with a migratory background enrolled in Dutch primary schools has increased to 25% of the total Dutch primary school population. This number includes all second- and first-generation immigrants (Inspectie van het Onderwijs, 2019).

[2] See, e.g., Checchi and Flabbi (2006), Hanushek and Woessmann (2006), Brunello and Checchi (2007), Crul and Schneider (2010), Van de Werfhorst and Mijs (2010), Crul et al. (2012), Bol and van de Werfhorst (2013), Ludemann and Schwerdt (2013), Hopwood et al. (2016).

[3] Enrollment level refers to the secondary school level a student enrolled in. So it does not refer to the number of enrolled students.

[4] Children enrolled in special education are not part of our study. These children need extra help with learning and other development issues and get more personalized guidance, in smaller classes, at a special education school. They do not take the primary school exit test at the end of primary school.

Using a difference-in-differences methodology, we compare the secondary school enrollment levels of migrants to those of native students. From these analyses, we draw two main conclusions. Firstly, the event studies show the commonality of the trends in enrollment levels between the groups before the policy change implying that we can use the DiD methodology. Secondly, the main result of the analyses is that the secondary school enrollment level of migrants increased substantially relative to that of natives due to the policy change.

An interesting question is which mechanism induced changes in enrollment level. We analyze whether the policy change affected school exit tests and teacher track recommendations as potential channels. Our results show relative increases for migrants both in exit test scores and track recommendations. In addition, after the policy change, teachers could upgrade their initial track recommendation after the primary school exit test result became available. We find a relative increase for migrants of the revised track recommendation as well. Therefore, teachers seem to give higher initial advice, and after seeing the higher relative test scores of migrants, further revise their track recommendations of migrants upward.

Earlier literature regarding the same policy implementation shows that students are more likely to change to a different secondary education track than recommended by the teacher when the teacher no longer uses the test score while giving the secondary school track recommendation (Swart et al., 2019). Feron (2018) shows that teachers gave more favorable assessments to girls and students with highly educated parents after the policy change. The policy change has not been explicitly linked to migrant status, even though the student population with a migrant background is growing in the Netherlands. This paper contributes to studies on the effects of this policy change by focusing on first generation migrants.

Our analysis of the effects on teacher track recommendations relates to research on teacher bias and statistical discrimination (Alesina et al., 2018; Burgess & Greaves, 2013). Driessen et al. (2008) mention that the sociologist literature on the case of the Netherlands has found that given comparable achievement, minority children are given a higher educational recommendation than non-minority children. Our analysis contributes to this literature by studying whether the policy change has affected this bias.

The postponement of the exit test relates to the literature on early tracking (e.g. Checchi & Flabbi, 2006; Hanushek & Woessmann, 2006; Brunello & Checchi, 2007; Bol & van de Werfhorst, 2013; Ludemann & Schwerdt, 2013; Hopwood et al., 2016) and the catching up process of migrant students (Evans & Fitzgerald, 2017; Gerritsen et al., 2019). The rich administrative data that we use allow us to study different mechanisms underlying the policy change.

The remainder of this paper is organized as follows. Section 2 provides an overview of the Dutch education system. Section 3 explains the policy change and describes the expected effects of this change. The empirical strategy is discussed in Sect. 4. Section 5 describes the data. Section 6 presents the findings. Section 7 concludes and discusses the results.

2 Dutch Education System

In the Netherlands, schooling is compulsory between the ages 5 to 16,[5] but usually students enter first grade when they are 4 years old. Primary school consists of 8 grades, including 2 years of kindergarten. The Netherlands has a relatively early tracking system. At the end of primary school, when students are on average 11 or 12 years old, they take a primary school exit test and receive a secondary school track recommendation by their teacher. There are eight secondary education track recommendations in three main education levels, varying in number of years of secondary schooling (from 4 to 6 years) and focus on vocational or general programs.[6] Each of these tracks has different tertiary education enrollment options. Figure 6 in the appendix displays an overview of the Dutch education system. Table 6 shows the different secondary education tracks and levels. In the analyses, we group the educational tracks to the three main Dutch secondary education school levels, which consist of pre-vocational (VMBO), secondary general (HAVO), and pre-university (VWO) education. We use the levels as a continuous variable (1–3).

Secondary school track placement is based on two observable signals about the ability of the child: the primary school exit test score and the secondary school track recommendation of the teacher. The primary school exit test takes place at the end of primary school. Students participate in a 2-day test which consists of multiple-choice questions on mathematics and Dutch language. Optionally, study skills and science can be tested as well, but the results on these topics are not taken into account in the calculation of the final test score. The average test scores are comparable across years, since the test scores are calibrated every year on a scale from 500 to 550. The answers are machine graded. Therefore, this primary school exit test score can be seen as an objective measure of a student's ability. Table 6 displays how these test scores can be converted to secondary track recommendations.[7]

The teachers' secondary school track recommendations, on the other hand, are a more subjective measure for student ability. Teachers can use several sources of

[5] When students do not have a secondary vocational, a secondary general or a pre-university degree before the age of 16, they have to stay in school until the age of 18.

[6] These tracks include: 2 tracks within pre-vocational education (vmbo-b and vmbo-k, 4 years), theoretical preparatory (vmbo-g(t), 4 years), secondary general, (havo, 5 years), and pre-university education (vwo, 6 years). Furthermore, primary school teachers can give double assessments, where they give a combination of two adjacent tracks as an assessment. There are three possibilities for this, as shown in Table 6. Besides the three main secondary education levels, the Netherlands also has a special education and a practical secondary education level. These two types of education levels are not included in this paper, since they are designed for students with low IQ or learning and behavioral difficulties. Due to the distinct learning abilities of the students that enroll at special and practical education, often they do not partake in the school exit test or have been enrolled at a special primary school. Therefore, there are no data available on their test score and teacher assessment (if students were enrolled in special primary education, it is very likely that they continue in special secondary education).

[7] The organization that creates the most used school exit test, Cito, suggests a secondary education track placement for each Cito test score (on a scale of 500–550). The Dutch Education Inspectorate uses a similar method to transfer school exit test scores into suggested track placements. Table 6 shows the method of the Dutch Education Inspectorate. We use this method to scale teacher assessments and test scores on a 3-point scale, since this method is most widely used by teachers throughout the Netherlands.

information in their decision process, such as previous test scores, experiences of other teachers, learning attitude and motivation of the student, gender, socio-economic status (SES), or ethnicity. There is no strong incentive for teachers to give assessments that are as high as possible, since their compensation scheme does not depend on the level of their assessments. Additionally, Feron (2018) points out that there is a repeated interaction between primary schools and the secondary schools students enroll at, which aligns the incentives and increases the understanding of the ability signals between primary school teachers and secondary schools over time.

3 Policy Change and Hypotheses

From the school year 2014–2015 onwards, there was a change in the Netherlands in the procedure for secondary school track assignment of students at the end of primary education. The main change in the procedure was that the exit test was delayed by three months, which in turn changed the relative timing of the teacher track recommendation and exit test.

Before the policy change, students would first take the school exit test, so that when teachers made their track recommendations, they took these school exit test scores into account. After the policy change, from the school year 2014–2015 onwards, teachers give their secondary school track recommendations before the students take the school exit test. We call these teacher track recommendations the *initial* teacher track recommendations. Once the school exit test scores are known, teachers can change their track recommendation upwards if a student has a higher school exit test score than the initial teacher track recommendation, although teachers are not obliged to do so. We call these teacher track recommendations after the test score is known *revised* teacher track recommendations. A detailed overview of the pre- and post-policy timing of events is shown in Fig. 1.

Hence, before the policy change, the secondary school track assignment largely depended on the school exit test score, which is an objective ability measure. After the policy change, the track assignment depended more on the track recommendation by the teacher(s), which is a subjective ability measure.[8]

The motivation behind the policy change to delay the test was threefold. First, by delaying the test, the teacher can no longer observe the test score while giving an assessment. Arguably, this results in the teacher using information on the development of the student over a longer time trajectory instead of a one-time ability measure as reported by the test score. Second, after the policy change, due to the postponement of the exit test, students who did not yet receive their preferred recommendation are forced to remain attentive in school for a longer time period instead of spending time on non-educational activities. The assumption here was

[8] The difference between scores on objective and subjective ability measures has been related to several background characteristics of students and to discrimination (Burgess & Greaves, 2013; De Boer et al., 2010; Dee, 2005; Feron, 2018; Gibbons & Chevalier, 2008; Glock et al., 2015; Lavy, 2008; Pit-ten Cate et al., 2016; Timmermans et al., 2015).

that once students had taken the school exit test, they would not pay much attention in class anymore, since what they did after the test did not count towards the track assignment decision. Third, the time between the exit test and the start of secondary education is shorter. This could lead to less skill deterioration and a more accurate ability measure for secondary education institutions (Rijksoverheid, 2016a; Feron 2018).

3.1 Expected Effects of the Policy Change

The policy change can have different effects on different groups of students. We expect different effects of the policy change on secondary-school enrollment levels for first-generation migrant students compared to the natives, because of differential effects on test scores and teacher track recommendations.

First, with the delay in the test moment, the test takes place after the teachers have given their initial track recommendations. In other words, teachers make an assessment without being able to observe the exit test scores. Following the literature on objective (blind testing such as the school exit test) and subjective (such as the teacher track recommendation) measures of ability (e.g., Alesina et al., 2018; Botelho et al., 2015; Burgess & Greaves, 2013; Hanna & Linden, 2012; Van Ewijk, 2011), we can form two conflicting hypotheses. The first is that the initial teacher track recommendation will be closer to the true ability of a migrant student without a school exit test score, because the test scores are relatively low for migrant students due to language barriers (Crawford, 2004). This is in line with one of the motivations for introducing the policy: the teacher assessment of migrant children is higher when a teacher cannot observe the school exit test score. Therefore, the share of children with a migration background who receive a higher track placement based on their initial teacher track recommendation should *increase* compared to the share of natives who receive a higher initial track recommendation after the policy change.

The second (conflicting) potential effect is that teacher track recommendations are less reflective of the migrant student's ability without a school exit test score because of teacher bias and statistical discrimination of migrant students. When teachers have less information on ability measures of students, they may use information on the mean score of students, e.g. of students with the same gender or ethnicity. Burgess and Greaves (2013) show that teachers statistically discriminate with respect to migration background for migrants in the United Kingdom, and Alesina et al. (2018) find evidence for this relationship for Italian teachers. This conflicting idea suggests that the share of children with a migration background who receive a higher initial teacher track recommendation should *decrease* compared to the share of natives who receive a higher initial teacher track recommendation after the policy implementation.

Secondly, after the policy implementation, the school exit test is taken three months later than before the implementation. We assume that first-generation migrant students are still learning Dutch and have a steeper learning curve in this catching up process. Hence, first-generation migrants may benefit from a

three-month delay in the exit test resulting in higher test scores for these migrant students.

Importantly for our analysis of the mechanisms, there were other changes which occurred at the point in time of the policy change. We discuss these in relation to our analysis of mechanisms in the section below.

4 Empirical Strategy

The empirical strategy relies on a difference-in-differences (DiD) approach in which secondary school enrollments of first-generation migrant students are compared to those of natives, before and after the policy change. The basic specification is as follows:

$$\begin{aligned} Y_i = {} & \alpha_0 + \alpha_1[Migrant_i \times Policy_i] + \alpha_2 Policy_i + \alpha_3 Migrant_i \\ & + \alpha_4 X_i + \alpha_5[Migrant_i \times 2011/12] + \alpha_6 2011/12 \\ & + \alpha_7[Migrant_i \times 2012/13] + \alpha_8 2012/13 + \varepsilon_i \end{aligned}$$

where the outcome variable (Y_i) is the secondary school level of enrollment in our main analysis, and in the analyses of the mechanisms, Y_i is either the test score or the initial/revised teacher track recommendation. $Migrant_i$ is a dummy which takes on the value 1 if the student has a first-generation migration background and 0 if the student is a native. $Policy_i$ is a dummy which is 1 in the school years after the policy change, X_i includes dummies for gender, age when taking the test, and age at arrival (for first-generation immigrants), and the equation includes an error term ε_i. The baseline year is the school year 2013/14 (i.e., the year prior to the policy implementation).

We are interested in the DiD estimator α_1, i.e. the interaction between the policy and first-generation migrant dummy. This estimator reports how the policy change affects the enrollment level in secondary school, and in the analyses of the mechanisms, it shows the effects on teacher assessments or exit test scores for first-generation immigrants compared to natives.

The identifying assumption in a DiD framework is that the trends in the outcome variables for the first-generation migrant group and natives would remain the same in absence of the policy change. Common pre-policy trends in treatment and control group suggest that the trends would also remain the same after the trend if the policy had not been implemented. To check whether the common trends assumption holds, we conduct event study analyses by including the interaction between $Migrant_i$ and pre-policy school years 2011/12, and 2012/13. The estimates reveal the difference between the groups in the pre-policy period. The common trends assumption holds if the pre-policy interactions are not significantly different from zero.[9]

[9] Before 2010- '11, the data were not filled well enough to be included in the analysis. Based on signals of low data quality of other users of the data at the time of analysis, we also excluded 2010- '11. However, while revising the paper for publication, we learned that the 2010- '11 data should have been of sufficient quality to be used in our analysis. Unfortunately, at that point in time, we no longer had access to the data. Future users of the data could include this year in their analysis.

We analyze the effects on enrollment into secondary school as our main analysis. Enrollment is measured in the third year of secondary school. We measure enrollment in this year instead of the first year of secondary school to make sure we capture structural enrollment of the student and not temporary enrollment in a level that is too high or low, and because in some schools the first two years contain mixed levels of secondary school.

The analyses in the mechanism section use exit test scores and teacher recommendations as outcome variables. There are several reasons why these analyses can only serve as indications of mechanisms and need to be interpreted with some caution. Firstly, the stakes of the test changed before and after the policy change. Before the policy change, the test was high stakes for all students, while afterwards, the test could only increase scores upward. This implies that it only was a high stakes test for those at the verge of reaching a higher secondary school level. Secondly, the test became compulsory after the policy change whereas it was not compulsory before the change (although almost all students participated in the test). Thirdly, the market for tests opened up after the policy change. Before the change, only CITO took tests, whereas after the change, other test takers also entered the market. To assess whether the second and third points are relevant concerns, our main analyses only use schools which took the CITO test before and after the change. In robustness checks, we show the results for all schools. Fourthly, as we will show below, test scores and teacher advice did not always have a common trend before the policy change.

5 Data

This paper uses Dutch administrative data that include all primary school students in the school years 2011/12 until 2017/18. The children are followed into their third year in secondary school, i.e. 2014/15 until 2021/22. The data include information on third year secondary school enrollment level, exit test scores[10] and teacher track recommendations. Moreover, the data contain several demographic variables (e.g., gender, age, country of origin, and arrival date).

We define the following groups. The *natives* are students for whom both parents are Dutch (77 percent of the students).[11] The second group consists of *first generation migrants*: i.e., students who were born abroad and who have at least one parent who was born abroad. The group includes first-generation Western immigrants (1 percent of the students) and the first-generation non-Western immigrants (1 percent

[10] In the analyses of the mechanisms, we only include students for whom we know the Cito test score. After the policy change, some students participated in a different exit school test, but these results are not comparable to the Cito results. Furthermore, the vast majority of the students who took a test, participated in the Cito test.

[11] The percentages reported in the text are calculated based on the selection of students for whom both a test score and teacher assessment are known.

Table 1 Descriptive statistics

Variable	Labels	Natives			1st generaton		
		Mean	sd	N	Mean	sd	N
Secondary school level	1: vmbo, 2: havo, 3: vwo	1.774	0.836	675,879	1.662	0.826	19,507
Test score	1: vmbo, 2: havo, 3: vwo	2.079	0.858	675,879	1.856	0.861	19,507
Teacher assessment	1: vmbo, 2: havo, 3: vwo	1.885	0.841	675,879	1.719	0.827	19,507
Revised teacher assessment	1: vmbo, 2: havo, 3: vwo	2.033	0.778	18,724	1.908	0.788	947
Male	0: female, 1: male	0.498	0.500	675,876	0.483	0.500	19,504
Age taking test		11.308	0.483	675,879	11.68	0.650	19,507
Socio-economic category parents	0: on benefits, 1: working	0.975	0.155	674,307	0.707	0.455	18,161
Income percentile parents		51.254	28.47	674,693	22.79	29.64	18,394
Indo-European language background	0: other language, 1: Indo-European				0.631	0.483	19,507

The statistics are based on students in the Netherlands who took the CITO test

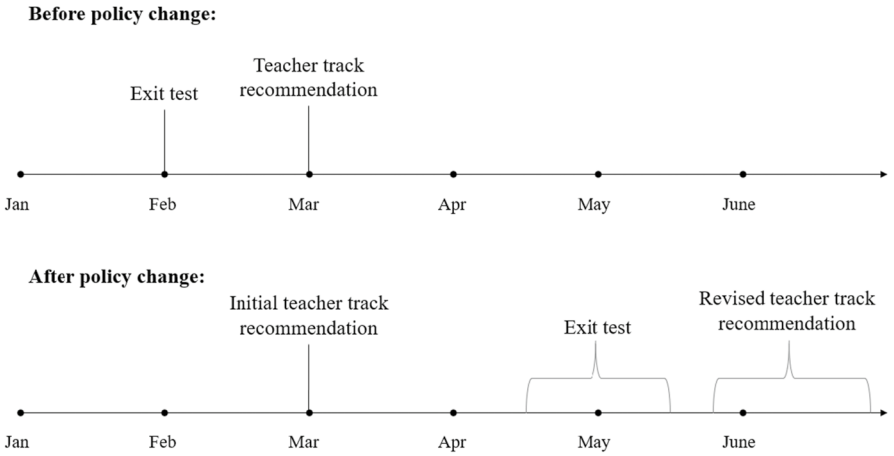

Fig. 1 Timing test score and teacher assessment pre- and post-policy change

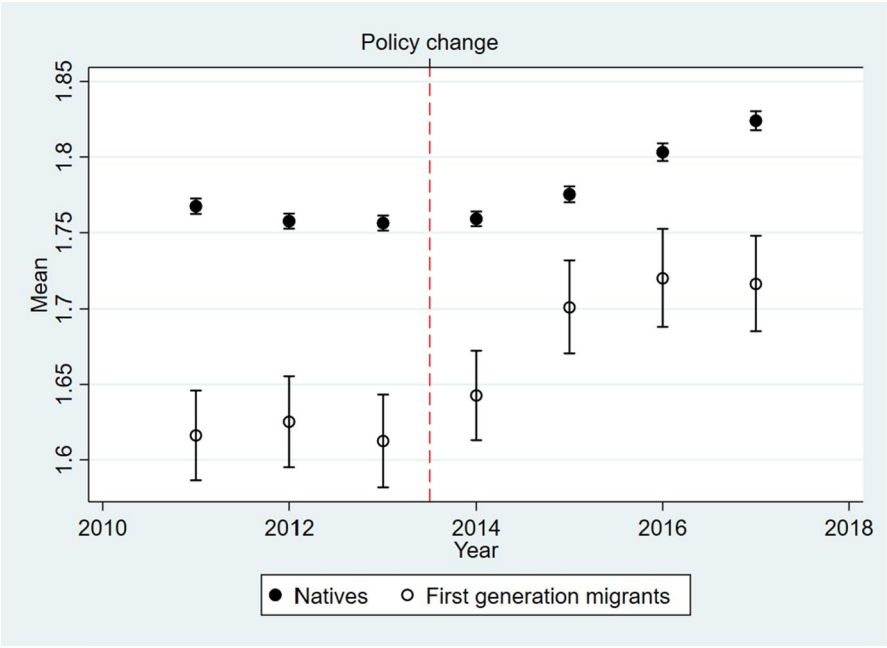

Fig. 2 Effect of the policy change on secondary school enrollment levels of migrants versus natives. *Note*: this figure shows the mean and 95% confidence interval of the secondary school level across years for first generation migrants and natives. The mean is based on a variable that has three values: (1) VMBO, (2) HAVO, (3) VWO. The graph is based on students in the Netherlands who took the CITO test

Table 2 Effect of the policy change on secondary school enrollment levels of migrants versus natives

	(1)	(2)	(3)	(4)
Migrant dummy	− 0.144***	0.221***	− 0.144***	0.222***
	(0.016)	(0.016)	(0.016)	(0.016)
2011 dummy	0.011***	0.017***	0.011***	0.017***
	(0.004)	(0.003)	(0.004)	(0.003)
2012 dummy	0.001	0.006*	0.001	0.006*
	(0.004)	(0.003)	(0.004)	(0.003)
2014 dummy	0.003	− 0.004		
	(0.004)	(0.003)		
2015 dummy	0.019***	0.007**		
	(0.004)	(0.003)		
2016 dummy	0.047***	0.029***		
	(0.004)	(0.004)		
2017 dummy	0.068***	0.043***		
	(0.004)	(0.004)		
Post-policy dummy (2014–17)			0.029***	0.015***
			(0.003)	(0.000)
2011 * migrant	− 0.007	− 0.001	− 0.007	− 0.001
	(0.023)	(0.022)	(0.023)	(0.022)
2012 * migrant	0.011	0.025	0.011	0.025
	(0.023)	(0.022)	(0.023)	(0.022)
2014 * migrant	0.027	0.032		
	(0.023)	(0.022)		
2015 * migrant	0.070***	0.084***		
	(0.023)	(0.022)		
2016 * migrant	0.061***	0.080***		
	(0.023)	(0.022)		
2017 * migrant	0.037	0.060***		
	(0.023)	(0.022)		
Post-policy * migrant			0.053***	0.067***
			(0.018)	(0.018)
Controls	No	Yes	No	Yes
Constant	1.756***	4.427***	1.756***	4.436***
	(0.003)	(0.023)	(0.003)	(0.023)
Observations	695,386	692,468	695,386	692,468
R-squared	0.001	0.139	0.001	0.139

The dependent variable is the secondary school enrollment level, measured in the third year of high school, on the scale: (1) VMBO, (2) HAVO and (3) VWO. The table displays results for schools which used the CITO test throughout the years of analysis. Columns 2 and 4 control for male, age when taking the test, income percentile parents, and socio-economic category parents

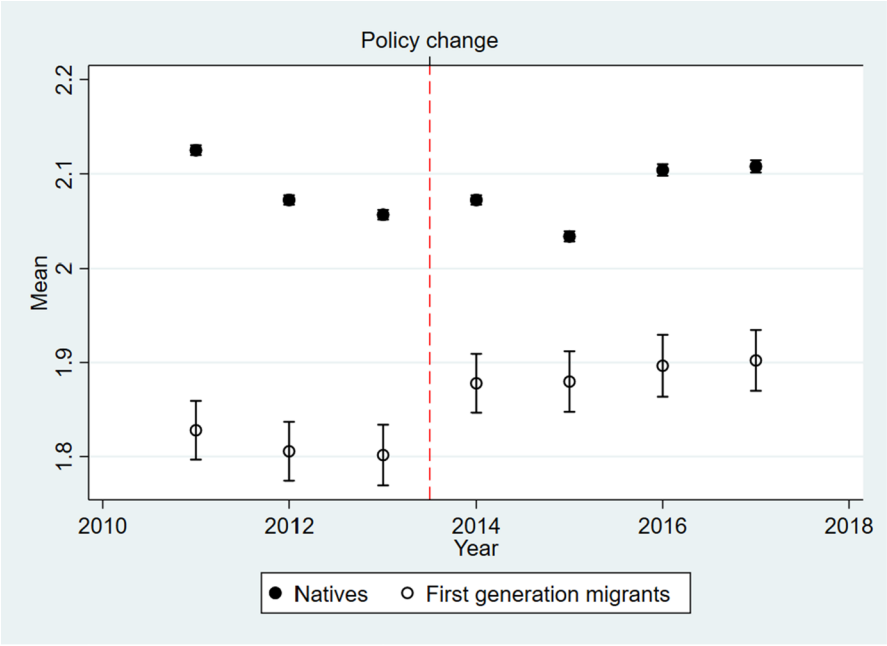

Fig. 3 Effect of the policy change on primary school exit test scores of migrants versus natives. *Note*: this figure shows the mean and 95% confidence interval of the secondary school level based on primary school exit test score thresholds, across years for first generation migrants and natives. The mean is based on a variable that has three values: (1) VMBO, (2) HAVO, (3) VWO. The graph is based on students in the Netherlands who took the CITO test

of the students).[12] A third group, *second generation migrants,* are also a part of the total student population (21%) but are not included in our analyses. Table 1 provides summary statistics of the data (across all years) which show that migrants have lower enrollment levels, lower teacher advice, and lower exit test scores. This table shows statistics only including schools that took the CITO test throughout the period of analysis. Table 7 shows similar statistics for the full population.

[12] First-generation Western immigrants and their parent(s) are born in Europe (excluding Turkey), North-America, Oceania, Japan or Indonesia. First-generation non-Western immigrants and their parent(s) are born in Africa, South-America, Asia or Turkey.

Table 3 Effect of the policy change on primary school exit test scores of migrants versus natives

	(1)	(2)	(3)	(4)
Migrant dummy	− 0.255***	0.075***	− 0.255***	0.075***
	(0.017)	(0.017)	(0.017)	(0.017)
2011 dummy	0.068***	0.075***	0.068***	0.075***
	(0.004)	(0.004)	(0.004)	(0.004)
2012 dummy	0.016***	0.021***	0.016***	0.021***
	(0.004)	(0.004)	(0.004)	(0.004)
2014 dummy	0.016***	0.009**		
	(0.004)	(0.003)		
2015 dummy	− 0.023***	− 0.035***		
	(0.004)	(0.004)		
2016 dummy	0.048***	0.030***		
	(0.004)	(0.004)		
2017 dummy	0.051***	0.025***		
	(0.004)	(0.004)		
Post-policy dummy (2014–17)			0.019***	0.005
			(0.003)	(0.003)
2011 * migrant	− 0.042*	− 0.043*	− 0.042*	− 0.043*
	(0.023)	(0.023)	(0.023)	(0.023)
2012 * migrant	− 0.012	− 0.001	− 0.012	0.000
	(0.024)	(0.023)	(0.024)	(0.023)
2014 * migrant	0.061***	0.061***		
	(0.023)	(0.023)		
2015 * migrant	0.101***	0.119***		
	(0.023)	(0.023)		
2016 * migrant	0.048**	0.062***		
	(0.024)	(0.024)		
2017 * migrant	0.050**	0.074***		
	(0.024)	(0.024)		
Post-policy * migrant			0.068***	0.081***
			(0.019)	(0.019)
Controls	No	Yes	No	Yes
Constant	2.056***	4.969***	2.056***	4.975***
	(0.003)	(0.024)	(0.003)	(0.024)
Observations	695,386	692,468	695,386	692,468
R-squared	0.003	0.092	0.002	0.091

The dependent variable is the primary school exit test score, measured on the scale: (1) VMBO, (2) HAVO and (3) VWO. The table displays results for schools which used the CITO test throughout the years of analysis. Columns 2 and 4 control for male, age when taking the test, income percentile parents, and socio-economic category parents

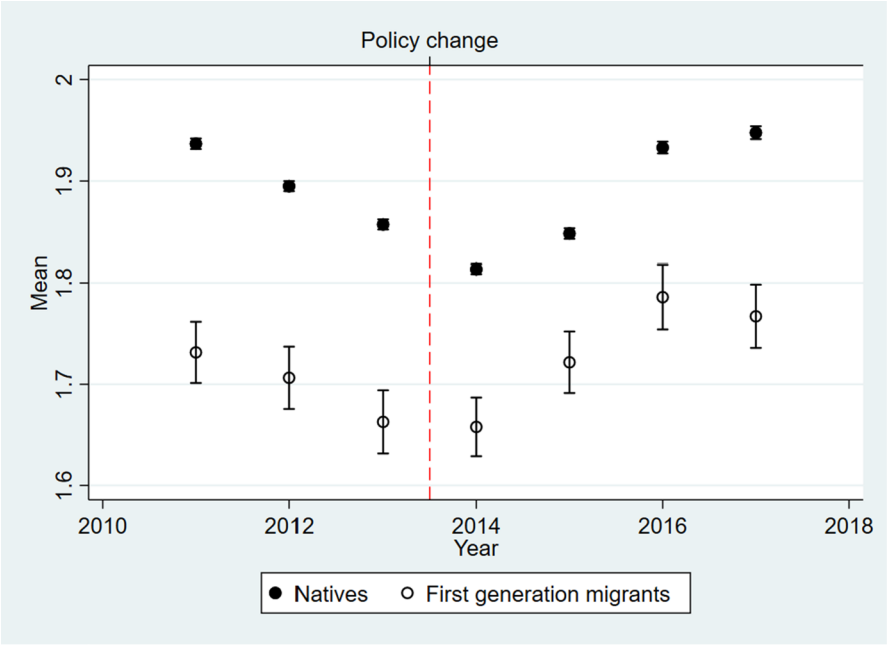

Fig. 4 Effect of the policy change on primary school teacher advice of migrants versus natives. *Note*: this figure shows the mean and 95% confidence interval of the secondary school level based on primary school teacher advice, across years for first generation migrants and natives. The mean is based on a variable that has three values: (1) VMBO, (2) HAVO, (3) VWO. The graph is based on students in the Netherlands who took the CITO test

6 Results

Figure 2 shows our main result: the enrollment level in secondary school for natives and migrants per year. The enrollment level of the natives remained relatively stable before the policy change, while it increased after the policy change. For migrants, the level was also constant before the policy change, and increased after the policy change. The increase for migrants is much more pronounced than the increase for the natives. As a result, the gap between natives and migrants decreases substantially.

This graph displays two points which we will analyze in more detail below. First, the trends in the enrollment levels seem common for both groups before the policy change. This is a crucial condition for using the difference-in-differences method which we will employ. Secondly, the effect of the policy change seems to have been more positive for first-generation students than for natives. Recall that after the policy change, schools could choose which exit test to take, while before the change, CITO was the sole provider of tests. Therefore, our baseline analysis includes schools which used the CITO test throughout the period. Graph A2 shows that our result remains robust if we analyze the trends for the full population.

Table 4 Effect of the policy change on primary school teacher advice of migrants versus natives

	(1)	(2)	(3)	(4)
Migrant dummy	− 0.194***	0.148***	− 0.194***	0.149***
	(0.017)	(0.016)	(0.017)	(0.016)
2011 dummy	0.079***	0.086***	0.079***	0.086***
	(0.004)	(0.003)	(0.004)	(0.003)
2012 dummy	0.038***	0.043***	0.038***	0.043***
	(0.004)	(0.003)	(0.004)	(0.003)
2014 dummy	− 0.044***	− 0.050***		
	(0.004)	(0.003)		
2015 dummy	− 0.009**	− 0.021***		
	(0.004)	(0.004)		
2016 dummy	0.076***	0.058***		
	(0.004)	(0.004)		
2017 dummy	0.090***	0.065***		
	(0.004)	(0.004)		
Post-policy dummy (2014–17)			0.018***	0.004
			(0.003)	(0.003)
2011 * migrant	− 0.011	− 0.007	− 0.011	− 0.007
	(0.023)	(0.022)	(0.023)	(0.022)
2012 * migrant	0.006	0.016	0.006	0.017
	(0.023)	(0.022)	(0.023)	(0.022)
2014 * migrant	0.039*	0.042*		
	(0.023)	(0.022)		
2015 * migrant	0.068***	0.085***		
	(0.023)	(0.022)		
2016 * migrant	0.048**	0.065***		
	(0.023)	(0.023)		
2017 * migrant	0.014	0.037		
	(0.023)	(0.023)		
Post-policy * migrant			0.051***	0.065***
			(0.018)	(0.018)
Controls	No	Yes	No	Yes
Constant	1.857***	4.632***	1.857***	4.655***
	(0.003)	(0.023)	(0.003)	(0.023)
Observations	695,386	692,468	695,386	692,468
R-squared	0.004	0.116	0.002	0.114

The dependent variable is the primary school teacher advice, measured on the scale: (1) VMBO, (2) HAVO and (3) VWO. The table displays results for schools which used the CITO test throughout the years of analysis. Columns 2 and 4 control for male, age when taking the test, income percentile parents, and socio-economic category parents

Table 2 reports our main analyses of the changes in secondary school enrollment levels of first-generation students and native students across time using regressions.

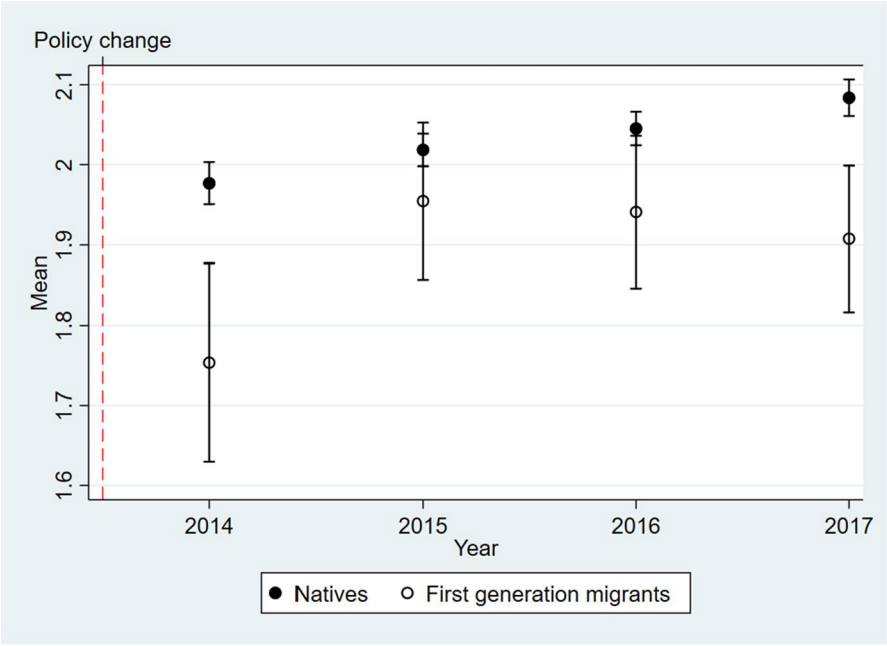

Fig. 5 Effect of the policy change on primary school revised teacher advice of migrants versus natives. *Note*: this figure shows the mean and 95% confidence interval of the secondary school level based on revised primary school exit test score thresholds, across years for first generation migrants and natives. The mean is based on a variable that has three values: (1) VMBO, (2) HAVO, (3) VWO. The graph is based on students in the Netherlands who took the CITO test

This allows us to control for various characteristics and analyze whether the common trend and the indicated changes also hold in a statistical sense. In this table, and all further tables, we will also show results for the schools which took the CITO test throughout the full period.

Firstly, the insignificant interactions before the policy change "2011*Migrant" and "2012*Migrant" confirm that the common trend holds in these years. This remains robust when controlling for covariates. These results indicate that we can apply the Difference-in-Differences technique to compare the development of secondary school levels for first-generation migrants to that of natives.

Secondly, the effect of the policy change across all years after the policy is given by the interaction "Post-policy*Migrant." This interaction shows that enrollment levels increased significantly for first-generation migrants compared to the increase for native students.

The relative increase is significant with an effect size of 0.053–0.067. One way of interpreting this result is that on top of the trend of natives, roughly 1 more

Table 5 Effect of the policy change on primary school revised teacher advice of migrants versus natives

	(1)	(2)	(3)	(4)
Migrant dummy	− 0.194***	0.149***	− 0.194***	0.150***
	(0.017)	(0.016)	(0.017)	(0.016)
2011 dummy	0.079***	0.086***	0.079***	0.086***
	(0.004)	(0.003)	(0.004)	(0.003)
2012 dummy	0.038***	0.043***	0.038***	0.043***
	(0.004)	(0.003)	(0.004)	(0.003)
2014 dummy	− 0.027***	− 0.034***		
	(0.004)	(0.003)		
2015 dummy	0.022***	0.010***		
	(0.004)	(0.004)		
2016 dummy	0.108***	0.090***		
	(0.004)	(0.004)		
2017 dummy	0.121***	0.095***		
	(0.004)	(0.004)		
Post-policy dummy (2014–17)			0.045***	0.030***
			(0.003)	(0.003)
2011 * migrant	− 0.011	− 0.007	− 0.011	− 0.007
	(0.023)	(0.022)	(0.023)	(0.022)
2012 * migrant	0.006	0.017	0.006	0.017
	(0.023)	(0.022)	(0.023)	(0.023)
2014 * migrant	0.042*	0.045**		
	(0.023)	(0.022)		
2015 * migrant	0.080***	0.098***		
	(0.023)	(0.022)		
2016 * migrant	0.060**	0.078***		
	(0.023)	(0.023)		
2017 * migrant	0.031	0.055**		
	(0.023)	(0.023)		
Post-policy * migrant			0.063***	0.077***
			(0.018)	(0.018)
Controls	No	Yes	No	Yes
Constant	1.857***	4.668***	1.857***	4.695***
	(0.003)	(0.024)	(0.003)	(0.024)
Observations	695,386	692,468	695,386	692,468
R-squared	0.005	0.117	0.002	0.114

The dependent variable is the primary school teacher advice, measured on the scale: (1) VMBO, (2) HAVO and (3) VWO. The table displays results for schools which used the CITO test throughout the years of analysis. The common trend is based on the initial teacher advice. Columns 2 and 4 control for male, age when taking the test, income percentile parents, and socio-economic category parents

migrant out of 15–20 migrants will enroll in a higher-level secondary school level. This is very large for merely three months additional learning time.[13]

6.1 Mechanisms

In this section, we investigate underlying reasons for the relative change in enrollment level. We analyze effects on primary school exit test scores, and on initial and revised teacher recommendations.

Figure 3 reports the mean exit test scores for each group per year. The test scores are originally on a scale from 500 to 550, but recoded to the three main levels of education (1 VMBO, 2 HAVO, 3 VWO). Native students score higher than migrants. In the year after the policy change, the test scores of native students remained similar. Those of first-generation immigrants increased after the policy change. Figure 8 shows that this result is again robust when we use all schools in the analysis. Therefore, the policy change appears to have had a positive relative effect on migrants' exit test scores.

Table 3 confirms these findings in a statistical sense. However, note that this table also reveals that the common trend assumption does not hold in the year 2011–12. This implies that we need to interpret the results with caution.

Figure 4 reports the average initial teacher track recommendations for the groups across years. In line with the findings above, initial teacher track recommendations are highest for the natives and lowest for the first-generation students. After the policy change, there is a decrease in initial teacher track recommendations for natives, but we can see a stable pattern for migrants. This indicates that the relative advice for migrants improved. Figure 9 shows that the result remains robust if we use all schools. Table 4 shows the corresponding regression analyses of these analyses. The trends are common.

Figure 5, 10, and Table 5 show the results for the revised teacher advice. The common trend in Table 5 is not displayed in the figure because there was no revised teacher advice possible before the policy change. In the table, the common trend is based on the initial teacher advice. Recall that the actual secondary school track placement decision after the policy change is based on the revised teacher track recommendation. When a child obtains a higher school exit test score than the initial teacher track recommendation, teachers can change their track recommendation upwards. The policy affected the revised teacher track recommendation of natives and migrants. This result is in line with Swart et al. (2019). The change from initial to revised teacher track recommendations is larger for migrants than for natives, indicating that teachers responded to the relative increase in exit test scores.

In sum, while the results from our main analysis show a clear common trend in the full period before the policy change, the trend appears less common in some of

[13] In further (unreported) analyses, we distinguished migrants with an Indo-European language background from those without one. We find similar effects in both groups with respect to all outcome variables.

the analyses of the mechanisms. This implies that we can use a DiD method for our main variable enrollment level in secondary school, but we need to be more cautious using this technique in the mechanism section.

7 Conclusions

This paper investigates the effects of a policy change regarding tracking into secondary school on the secondary school enrollment levels of first-generation migrants compared to those of natives. The policy change entails that the primary school exit test was delayed by three months and, consequently, was held after the initial teachers' track recommendations. Students, therefore, received (a) an initial recommendation based on other factors than the exit test scores, (b) 3 months extra to learn for the exit test, and (c) the possibility to have their teacher recommendation adjusted based on the exit test scores.

Our main results indicate that the policy change had a positive effect for first-generation migrant students: it increased the enrollment level of first-generation migrants relative to natives in secondary education. The gap between migrants and natives reduced by more than half. The effect size of the policy change is such that on top of the trend of natives, around 1 more migrant in every 15–20 migrants will attend a higher-level secondary school than before the change. This is a very large effect given the small treatment. Our analyses suggest that likely mechanisms for this effect are the increases of exit test scores and teacher assessments for migrants relative to natives.

Appendix

See Tables (6, 7) and Figs. (6, 7, 8 ,9 and 10).8,

Table 6 Teacher assessments and secondary education tracks

Secondary education track placement suggestion	Dutch abbreviation	ISCED level	Analysis
Pre-vocational	Vmbo-b	2	1
	Vmbo-b—Vmbo-k	2	1
	Vmbo-k	2	1
Theoretical preparatory	Vmbo-g(t)	2	1
Theoretical preparatory—Secondary General	Vmbo-g(t)—Havo	2–3	1
Secondary General	Havo	3	2
Secondary General—Pre-University	Havo—Vwo	3–4	2
Pre-University	Vwo	4	3

An overview of the Dutch secondary education tracks. For comparability reasons, the corresponding ISCED levels are displayed in column 2 and the Dutch abbreviations in column 3. The last column shows that practical education corresponds to the lowest track and pre-university to the highest track

Table 7 Descriptive statistics, full population

Variable	Labels	Natives			1st generaton		
		Mean	sd	N	Mean	sd	N
Secondary school level	1: vmbo, 2: havo, 3: vwo	1.768	0.834	798,144	1.659	0.823	24,152
Test score	1: vmbo, 2: havo, 3: vwo	2.056	0.856	798,144	1.839	0.855	24,152
Teacher assessment	1: vmbo, 2: havo, 3: vwo	1.880	0.840	798,144	1.714	0.825	24,152
Revised teacher assessment	1: vmbo, 2: havo, 3: vwo	2.013	0.777	27,393	1.869	0.778	1471
Male	0: female, 1: male	0.498	0.500	798,139	0.483	0.500	24,148
Age taking test		11.304	0.480	798,144	11.68	0.651	24,152
Socio-economic category parents	0: on benefits, 1: working	0.975	0.157	796,269	0.701	0.458	22,449
Income percentile parents		51.298	28.45	796,725	22.47	29.33	22,736
Indo-European language background	0: other language, 1: Indo-European				0.628	0.483	24,152

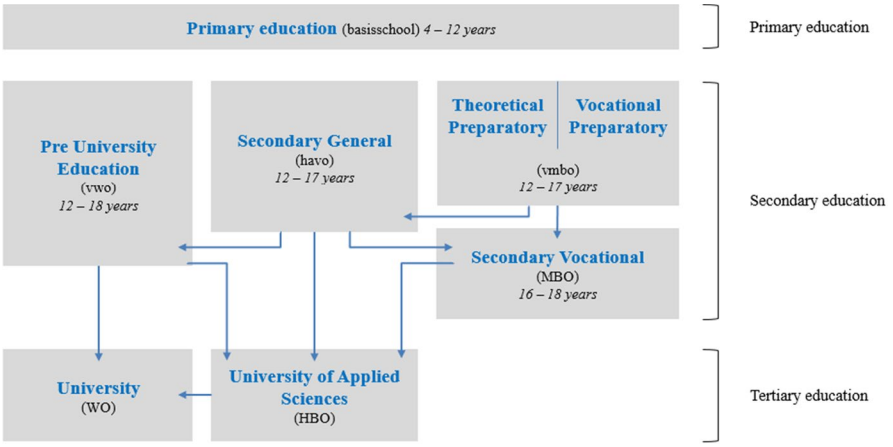

Fig. 6 Dutch education system. *Note*: This figure provides an overview of the Dutch education system. Special education and practical education are not included, since they are not considered to be part of mainstream education

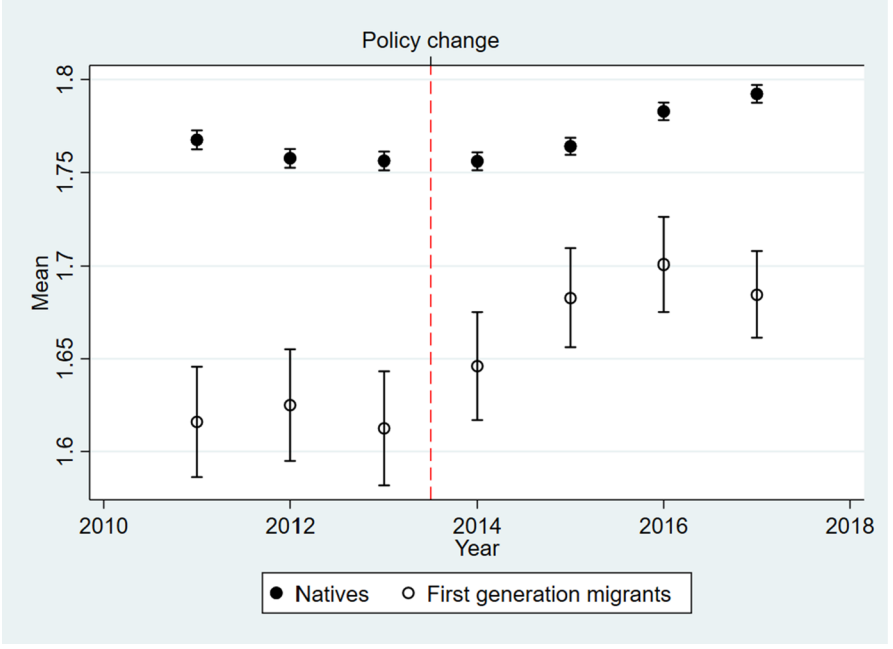

Fig. 7 Effect of the policy change on secondary school enrollment levels of migrants versus natives, full population. *Note*: this figure shows the mean and 95% confidence interval of the secondary school level across years for first generation migrants and natives. The mean is based on a variable that has three values: (1) VMBO, (2) HAVO, (3) VWO. The graph is based on all students in the Netherlands

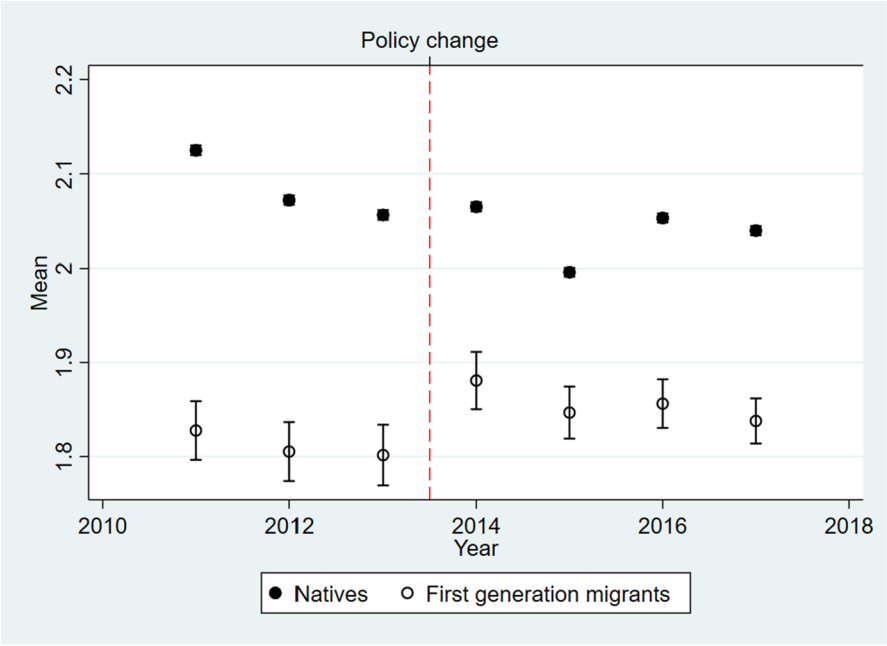

Fig. 8 Effect of the policy change on primary school exit test scores of migrants versus natives, full population. *Note*: this figure shows the mean and 95% confidence interval of the secondary school level based on primary school exit test score thresholds, across years for first generation migrants and natives. The mean is based on a variable that has three values: (1) VMBO, (2) HAVO, (3) VWO. The graph is based on all students in the Netherlands

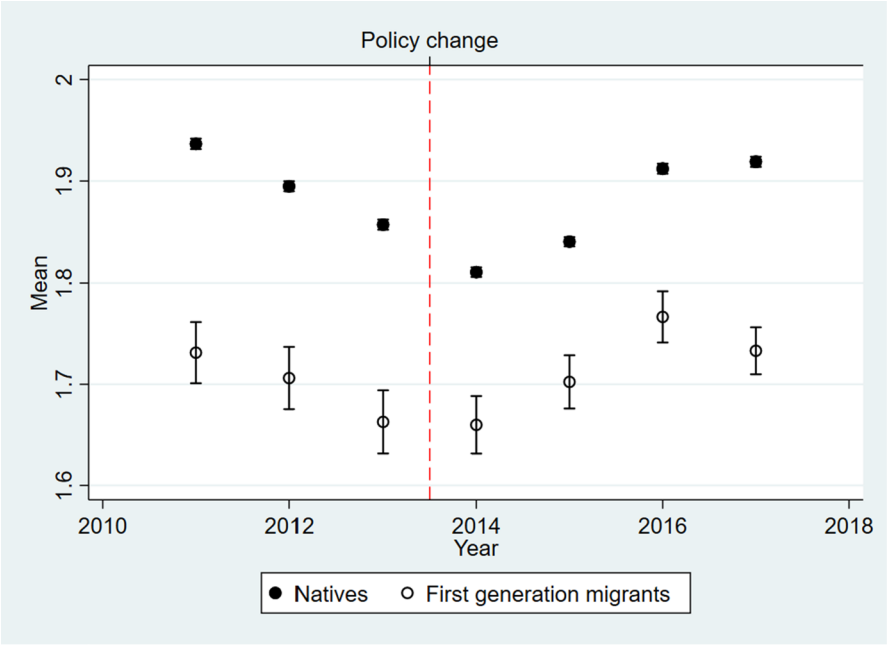

Fig. 9 Effect of the policy change on primary school teacher advice of migrants versus natives, full population. *Note*: this figure shows the mean and 95% confidence interval of the secondary school level based on primary school teacher advice thresholds, across years for first generation migrants and natives. The mean is based on a variable that has three values: (1) VMBO, (2) HAVO, (3) VWO. The graph is based on all students in the Netherlands

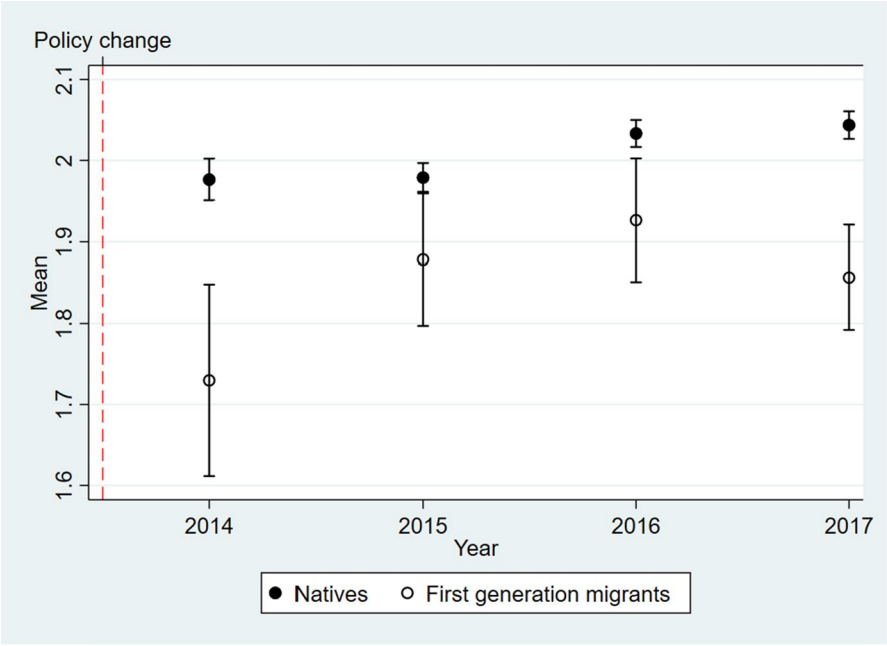

Fig. 10 Effect of the policy change on primary school revised teacher advice of migrants versus natives, full population. *Note*: this figure shows the mean and 95% confidence interval of the secondary school level based on revised primary school exit test score thresholds, across years for first generation migrants and natives. The mean is based on a variable that has three values: (1) VMBO, (2) HAVO, (3) VWO. The graph is based on all students in the Netherlands

Acknowledgements The authors received valuable comments from two referees, Ilja Cornelisz, Carla Haelermans, Tim Huijts, Chris van Klaveren, Arjan Non, Melline Somers, Stan Vermeulen, seminar participants at Sussex university, the Centre for Research and Analysis of Migration (London), Aachen University, the University of Bristol, and conference participants of the Nederlandse Economen Dag 2019, EALE/SOLE/AASLE 2020, IWAEE 2022, the 2023 Quality in Education conference at Stockholm School of Economics, the 2023 Koninklijke Vereniging voor de Staatshuishoudkunde meeting in The Hague, the 2023 Research Centre for Education and the Labour Market (ROA, Maastricht) conference and the 2024 Lisbon Economics and Statistics of Education conference.

Funding No funds, grants, or other support was received.

Declarations

Conflict of interest There are no conflicts of interests.

References

Alesina, A., Carlana, M., La Ferrara, E., & Pinotti, P. (2018). Revealing stereotypes: Evidence from immigrants in schools. NBER Working Paper No. 25333, Cambridge.

Bol, T., & Van de Werfhorst, H. G. (2013). Educational systems and the trade-off between labor market allocation and equality of educational opportunity. *Comparative Education Review, 57*(2), 285–308.

Botelho, F., Madeira, R. A., & Rangel, M. A. (2015). Racial discrimination in grading: Evidence from Brazil. *American Economic Journal: Applied Economics, 7*(4), 37–52.

Brunello, G., & Checchi, D. (2007). Does school tracking affect equality of opportunity? New international evidence. *Economic Policy, 22*, 781–861.

Burgess, S., & Greaves, E. (2013). Test scores, subjective assessment, and stereotyping of ethnic minorities. *Journal of Labor Economics, 31*, 535–576.

Checchi, D. & Flabbi, L. (2006). *Intergenerational mobility and schooling decisions in Italy and Germany*. IZA Discussion Paper 2876, IZA, Bonn.

Crawford, J. (2004). *No child left behind: Misguided approach to school accountability for English language learners*. Forum on ideas to improve the NCLB accountability provisions for students with disabilities and English language learners. Center on Education Policy, Washington D.C.

Crul, M., & Schneider, J. (2010). Comparative integration context theory: Participation and belonging in new diverse european cities. *Ethnic and Racial Studies, 33*(7), 1249–1268.

Crul, M., Schneider, J., & Lelie, F. (2012). *The European second generation compared: Does the integration context matter?* Amsterdam University Press.

De Boer, H., Bosker, R. J., & van der Werf, M. P. C. (2010). Sustainability of teacher expectation bias effects on long-term student performance. *Journal of Educational Psychology, 102*, 168–179.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review, 95*(2), 158–165.

Driessen, G., Sleegers, P., & Smit, F. (2008). The transition from primary to secondary education: Meritocracy and ethnicity. *European Sociological Review, 24*(4), 527–542.

Evans, W. & Fitzgerald, D. (2017). *The economic and social outcomes of first-generation migrants in the United States: Evidence from the ACS*. NBER Working Paper No. 23498, National Bureau of Economic Research, Cambridge.

Feron, E. (2018). *The role of cognitive tests and teachers in the transition from primary to secondary education*. Datawyse/Universitaire Pers Maastricht.

Gerritsen, S., Kattenberg, M., & Kuijpers, S. (2019). *The impact of age at arrival on education and mental health*. CPB Discussion paper, CPB, Den Haag.

Gibbons, S., & Chevalier, A. (2008). Assessment and age 16+ education participation. *Research Papers in Education, 23*(2), 113–123.

Glock, S., Krolak-Schwerdt, S., & Pit-ten Cate, I. M. (2015). Are school placement recommendations accurate? The effect of students' ethnicity on teachers' judgments and recognition memory. *European Journal of Psychology of Education, 30*, 169–188.

Hanna, R. N., & Linden, L. L. (2012). Discrimination in grading. *American Economic Journal: Economic Policy, 4*(4), 146–168.

Hanushek, E. A., & Woessmann, L. (2006). Does educational tracking affect performance and inequality? Differences in differences evidence across countries. *The Economic Journal, 116*, C63–C76.

Hopwood, B., Hay, I., & Dyment, J. (2016). The transition from primary to secondary school: Teachers' perspectives. *The Australian Educational Researcher, 43*, 289–307.

Inspectie van het Onderwijs. (2018). *De staat van het Onderwijs 2018*. Inspectie van het Onderwijs.

Inspectie van het Onderwijs. (2019). *De staat van het Onderwijs 2019*. Inspectie van het Onderwijs.

Lavy, V. (2008). Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment. *Journal of Public Economics, 92*(10–11), 2083–2105.

Ludemann, E., & Schwerdt, G. (2013). Migration background and educational tracking. Is there a double disadvantage for second-generation immigrants? *Journal of Population Economics, 26*, 455–481.

Pit-ten Cate, I. M., Krolak-Schwerdt, S., & Glock, S. (2016). Accuracy of teachers' tracking decisions: Short- and long-term effects of accountability. *European Journal of Psychology of Education, 31*, 225–243.

Rijksoverheid. (2016a). Toelating voortgezet onderwijs gebaseerd op schooladvies. https://www.rijksoverheid.nl/onderwerpen/toelating-middelbare-school/inhoud/toelating-voortgezetonderwijs-gebaseerd-op-schooladvies

Rijksoverheid. (2016b). Verplichte eindtoets basisonderwijs. https://www.rijksoverheid.nl/onderwerpen/toelating-middelbare-school/inhoud/verplichte-eindtoetsbasisonderwijs

Swart, L., Van den Berge, W., & Visser, D. (2019). De waarde van eindtoetsen in het primair onderwijs. CPB Policy Brief, Den Haag

Timmermans, A. C., Kuyper, H., & van der Werf, M. P. C. (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*, 459–478.

Van de Werfhorst, H. G., & Mijs, J. J. B. (2010). Achievement inequality and the institutional structure of educational systems: A comparative perspective. *Annual Review of Sociology, 36*, 407–442.

Van Ewijk, R. (2011). Same work, lower grade? Student ethnicity and teachers' subjective assessments. *Economics of Education Review, 30*(5), 1045–1058.

van het Onderwijs, I. (2016). *De staat van het Onderwijs 2016*. Inspectie van het Onderwijs.