The Hidden Divide: School Segregation of Teachers in the Netherlands

Rafiq Friperson Hessel Oosterbeek Bas van der Klaauw*

Abstract

We use Dutch register data to document the understudied phenomenon of teacher segregation. We show that teachers in primary and secondary schools in the four largest cities of the country – Amsterdam, Rotterdam, The Hague and Utrecht – are segregated in terms of their migration and social backgrounds. While segregation by social background is not much higher than what would be expected under random teacher-school assignment, segregation by migration background is substantial even after accounting for randomness. Relating schools' teacher composition to their student composition, we find in most cases that schools with a high proportion of teachers from a particular background tend to have a high proportion of students from that same background.

*This version: June 2023. Friperson: VU University Amsterdam, Department of Economics, (r.j.friperson@vu.nl); Oosterbeek: University of Amsterdam, School of Economics (h.oosterbeek@uva.nl); Van der Klaauw: VU University Amsterdam (b.vander.klaauw@vu.nl); Friperson acknowledges financial support from a Doctoral Grant for teachers from the Dutch Science Foundation (NWO).

1 Introduction

While school segregation of **students** is well-documented (cf. Reardon and Owens, 2014; Oosterbeek et al., 2021), this is not the case for school segregation of **teachers**. So far, teacher segregation has been studied for teachers working in a specific sector of schools (Fagernäs and Pelkonen, 2017) or is based on self-reported shares of teachers from different groups (Frankenberg, 2008).[1] This paper presents more comprehensive results based on register data that cover the entire population of teachers in primary and secondary schools in the Netherlands. We focus on segregation of teachers by their migration background and by their social background.

Different arguments point in opposite directions regarding the desirability of teacher segregation on these dimensions. The growing literature showing that students benefit from being taught by a teacher "like them" (Gershenson et al., 2021; Goldhaber et al., 2019; Villegas and Irvine, 2010; Grissom et al., 2017; Egalite et al., 2015; Gershenson et al., 2016; Prokic-Breuer and Vermeulen, 2020) suggests that school segregation of teachers should mirror the existing school segregation of students. At the same time, if desegregation of students is a policy goal then desegregation of teachers might be a way to achieve that insofar as student segregation is affected by teacher segregation.[2]

Our register data contain information from all primary and secondary school teachers from the school years 2009 until 2019. This includes information of the school where someone is teaching as well as of their migration and social backgrounds.[3] The register data also contain school level information of the migration background of students and of the education level of their parents.

We focus on school segregation in the four largest cities of the Netherlands: Amsterdam, Rotterdam, The Hague and Utrecht. The main reasons for this are that the student populations in these cities are (much) more diverse than elsewhere and that the shares

[1] Fagernäs and Pelkonen (2017) report segregation by gender and caste of public sector teachers in India. This gives an incomplete picture in the presence of a selective private school sector. Frankenberg (2008) presents results from a survey among 1,000 public school teachers in the US who report their own race as well as the share of peers who are white.

[2] Desegregation is considered desirable because it can reduce the achievement gap between students from different backgrounds (Billings et al., 2013), and may even have effects on integration more broadly (Rao, 2019).

[3] Because information about income and attained education level is only available from the late 1990s onwards, we measure the social background of teachers by the education level of their siblings (if any).

of teachers (and students) with a migration background are substantial compared to those in other parts of the country. We analyze data separately for primary schools and secondary schools.

To measure segregation we compute the dissimilarity index, which is a measure of evenness, as well as the interaction index, which is a measure of exposure. We also calculate what the segregation indices would have been in case teachers are randomly assigned to schools. This takes into account that some degree of segregation might be due to sampling variation (cf. Allen et al., 2015).

Our results show that in each of the four cities teachers are segregated in terms of their migration and social backgrounds. This is true in both primary schools and secondary schools. Segregation by migration background remains substantial even after correcting for sampling variation. Teacher segregation by social background is not much larger than what could be due to sampling variation. In a final step, we calculate Pearson correlation coefficients of schools' proportions of teachers and students with a particular background. In most cases, these correlations are quite substantial. For example, the correlation between secondary schools' proportion of students with a migration background and their proportion of teachers with a migration background is in none of the four cities smaller than 0.77. The noteworthy exception occurs in Rotterdam where the correlation of schools' shares of teachers and students from low-educated families is much lower than in the other cities.

The rest of this paper is organized as follows. The next section provides a brief description of the setting of our study; primary and secondary schools in the Netherlands. Section 3 introduces the segregation measures that we use. Section 4 describes the data sources and presents summary statistics. Section 5 presents the results and Section 6 concludes.

2 Context

Almost all schools in the Netherlands are publicly funded and run by autonomous school boards. School boards have freedom in the way they organize the teaching in schools. The Dutch Education Inspectorate monitors and reviews the quality of schools. The Ministry of Education is responsible for the quality of schools and sets standards for examinations and teaching staff (degrees). The Ministry funds school boards on a

per student basis.

We study teacher segregation in primary and secondary education. Primary education consists of eight years (two years in kindergarten and six regular grade levels). In sixth grade students take a nationwide test and are assigned by their primary school teachers to one of four ability tracks in secondary education: vocational-elementary, vocational-theory, college and university. The two vocational tracks last four years and give access to subsequent vocational programs. The college track takes five years and gives access to professional colleges (applied universities). The university track takes six years and gives access to university education. Some secondary schools only offer one track, others offer two or three tracks.

The Dutch system is characterized by free choice for teachers, schools and students. Teachers can work at the school they want provided that the school wants to hire them and they possess the required teaching credentials. Schools are free to decide which teachers to hire and to set the salary level within collective labor agreement standards. Students can also freely choose which school to attend, where in secondary education this is restricted by the assigned ability track, and possibly by schools' capacity constraints.

3 Measuring Segregation

We indicate to what extent there is teacher (and student) segregation by computing two widely used segregation measures: the dissimilarity index (D) which captures evenness, and the interaction index (I) which captures exposure (Frankel and Volij, 2011; Allen et al., 2015).

3.1 Evenness

Evenness measures the spatial distributions of different groups across schools. Segregation on this dimension is smallest when the proportion of each group in each school is the same as the proportion of each group in the total city population. The most widely used evenness measure is the dissimilarity index (D) (Massey and Denton, 1988):

$$D = \frac{1}{2} \sum_{i=1}^{n} \left| \frac{a_i}{A} - \frac{b_i}{B} \right| \tag{1}$$

where n is the number of schools in a city, a_i is the number of teachers (or students) from group a in school i and A is the total number of teachers from that group in the city, b_i is the number of teachers from group b in school i and B is the total number of teachers from that group in the city.

The value of D is often interpreted as the fraction of teachers that needs to switch schools to create an equal distribution across all schools. This is, however, only an accurate interpretation if the two groups are equally sized. Assume that there are only two schools (1 and 2), that $a_1 = 7$, $a_2 = 3$, so $A = 10$, $b_1 = 48$, $b_2 = 52$, so $B = 100$. This results in $D = 0.22$ ($= \frac{1}{2} \times (|0.7 - 0.48| + |0.3 - 0.52|)$). The numbers of teachers that have to be reassigned so that both schools have five teachers from group A and 50 teachers from group B is only four; two teachers from group A (20%) have to switch from school 1 to school 2, while two teachers from group B (2%) have to switch from school 2 to school 1.[4]

Allen et al. (2015) have drawn attention to the fact that even if teachers would be randomly assigned to schools, it is unlikely that all schools have exactly the same proportions of teachers from each group. This is comparable to sampling variation and is more relevant if the overall share of minority teachers is small and if the average number of teachers per school is not so large. Not taking this sampling variation into account leads to an overestimation of segregation. We correct for it by calculating segregation indices for a counterfactual situation where all teachers are randomly assigned to schools. We do this 100 times and then take the average of the 100 simulations as a measure of the segregation index that is caused by sampling variation (D_{noise}). We then determine the adjusted segregation index D_{adj} as the difference between the observed raw index and the part that can be attributed to sampling variation: $D_{adj} = D_{obs} - D_{noise}$.

3.2 Exposure

Exposure represents the degree of potential interaction between different groups at the school level. It captures a weighted average of the shares of the teachers that work in the same school and are from another group (Massey and Denton, 1988).

[4] D ranges from 0 (no segregation) to 1 (complete segregation). A value of D between 0 and 0.3 is considered as low segregation, a value between 0.3 and 0.6 as medium segregation, and a value above 0.6 as high segregation (Massey and Tannen, 2015).

Massey and Denton (1988) note that evenness and exposure measures are often correlated but are conceptually different. This is because exposure measures depend on the relative sizes of the groups being compared, whereas evenness measures are independent of relative group sizes. Migrant teachers can be evenly distributed across schools, but they may at the same time experience little exposure to non-migrant teachers, if the share of migrant teachers in the city teacher population is large. On the other hand, if migrant teachers are a relatively small proportion, they are likely to experience high levels of exposure to non-migrant teachers no matter what the pattern of evenness is.

We measure exposure with the interaction index (I), this index measures the degree of potential contact between groups at schools:

$$I = \sum_{i=1}^{n} \left(\frac{a_i}{A}\right) \left(\frac{b_i}{a_i + b_i}\right) \tag{2}$$

The value of I indicates the probability that a teacher from group A works at the same school as a teacher from group B. In the previous example with two schools, the value of I would equal 0.895 ($\approx \frac{7}{10} \cdot \frac{48}{55} + \frac{3}{10} \cdot \frac{52}{55}$). The minimum value, keeping school sizes fixed would be 0.818 ($\approx \frac{10}{10} \cdot \frac{45}{55} + \frac{0}{10} \cdot \frac{55}{55}$), while the maximum value is 0.909 ($\approx \frac{5}{10} \cdot \frac{50}{55} + \frac{5}{10} \cdot \frac{50}{55}$). These numbers reflect the average share of teachers from the majority group (B) that teachers from the minority group (A) meet in their school. Note that these numbers should be divided by 10 when we are interested in the average share of teachers from the minority group (A) that teachers from the majority group (B) meet in their school. This illustrates the dependence of I on the relative sizes of the groups.

4 Data and Summary Statistics

4.1 Data sources

We use administrative data from two sources: register data on teachers from the Education Executive Agency (DUO) of the Dutch Ministry of Education and register data on students and schools from Statistics Netherlands (CBS).

The teacher register data from DUO have information on all teachers who taught at primary or secondary public schools in any of the school years 2006/7 until 2019/20.[5]

[5]Schools are legally required to yearly report educational personnel to DUO. In the years 2009 to

For each teacher, the data contain the school address[6], type of job (temporary/temporary replacement or permanent) and job size (in full time equivalents). We enriched the DUO data in the microdata environment of Statistics Netherlands with information about teachers' migration background and the education level of their siblings.

The student register data from Statistics Netherlands have information on students who were enrolled in primary and secondary schools in the school years 2009/10 until 2019/20. For each student the data contain the school-location address, grade level and ability track. We enriched the student data with information about the students' migration background and education of their parents. Students and teachers with an unknown school(-location) are excluded from the analysis.

A student or teacher is labeled as having a migration background if at least one (grand)parent was born abroad. A migrant is labeled as 'non-western' if at least one (grand)parent originates from a country in Africa, South America or Asia (excluding Indonesia and Japan) or from Turkey. All other migrants are labeled as 'western'.[7] For the levels of education (lower, medium, high) we follow the definition of Statistics Netherlands.[8] We label that a student is from a low-educated family if its parents are low-educated. We label that a teacher is from a low-educated family if that teacher has a low-educated sibling.

We aggregate the data at the school level and at the school-location level to compute segregation indices over time separately for the four largest cities of the Netherlands.

2020 only a small percentage of educational personnel data was not collected by DUO (on average $< 1\%$ primary schools and $< 2\%$ secondary school teachers).

[6]Teacher data is collected at the school level (BRIN4), while student data is collected at the school-location level (BRIN6). In the school year 2019/20 primary schools in the four largest Dutch cities had on average 1.1 locations, and secondary schools in these cities had on average 2.2 locations, see Table 1.

[7]This is the definition of non-western from Statistics Netherlands. In the Spring of 2021 Statistics Netherlands stopped reporting migrants in terms of western and non-western origin. Statistics Netherlands did this because the Dutch Scientific Council for Government Policy (WRR) published an advice in which it argues that the distinction has no scientific foundation and has a colonial connotation.

[8]Lower education includes programs up to the lower secondary (vocational) level. Higher education refers to associate degree programs and above. Medium education refers to all other programs.

4.2 Summary Statistics

Table 1 shows the numbers of teachers (in full time equivalents), students, schools and locations by level of education (primary and secondary) for the school year 2019/20. This is reported for each of the four largest cities as well as for the rest of the country. The primary school teachers in the four largest cities represent approximately 15% of the total primary school teacher workforce in the Netherlands. For secondary school teachers this figure is 14%. The student-teacher ratio in primary schools is between 17.3 and 18.2 in the four largest cities and 19.1 in the rest of the country. In secondary schools the student-teacher ratio is lowest in Amsterdam (14.4) and Rotterdam (15.2) and is around 16.5 in The Hague, Utrecht and the rest of the country.

Primary schools in the four largest cities have on average more than 300 students, while average primary school size in the rest of the country is 230. For secondary schools the picture is reversed. In the largest cities average secondary school size is between 1140 (Utrecht) and 1500 (The Hague), while it is 1600 for secondary schools in the rest of the country.

Table 1: Numbers of teachers, students and schools by level and city; school year 2019-2020

A: Primary Schools	Teachers	Students	Schools	Locations
Amsterdam	3495	61919	197	210
Rotterdam	2969	52067	159	170
The Hague	2705	46823	132	135
Utrecht	1659	30184	91	98
Rest of the country	63248	1205581	5221	5683
B: Secondary Schools	Teachers	Students	Schools	Locations
Amsterdam	2950	42575	34	76
Rotterdam	2355	35716	29	71
The Hague	1764	28535	19	47
Utrecht	958	15921	14	21
Rest of the country	50620	827506	516	1235

Note: Teachers are measured in full-time equivalents.

Table 2 shows the means and standard deviations of schools' shares of teachers and students with a non-western migration background (panel A) and of teachers and students from low-education families (panel B) by level of education (primary and secondary) and city for school year 2019/20. In primary schools in the four largest cities, the mean share of teachers with a non-western migration background ranges from 0.07 in Utrecht to 0.18 in Amsterdam. In the rest of the country the share is only 0.02. The shares for teachers are much lower than for students. In primary schools in the largest cities, the mean share of students with a non-western migration background ranges from 0.37 in Utrecht to 0.58 in Rotterdam, while it is on average 0.18 in schools in the rest of the country. The shares of non-western teachers in secondary schools are somewhat higher than in primary schools, while for students these shares are somewhat lower than in primary schools.

The share of teachers from low-educated families is around 0.10 and this does not vary much between different locations and between primary and secondary schools. The share of students from low-educated families is on average between 0.20 and 0.25 in primary and secondary schools Amsterdam, Rotterdam and The Hague. These shares are somewhat lower in Utrecht and in the rest of the country.

Panel C of Table 2 reports the correlations between a school's share of teachers (students) with a non-western migration background and its share of teachers (students) coming from a low-education family by level of education and city. This shows that for students the two shares are very highly correlated, with correlation coefficients around 0.9. For teachers the correlations between the two shares are weaker, with correlation coefficients varying between 0.29 for secondary-school teachers in Rotterdam and 0.62 for secondary-school teachers in Utrecht.

Figure B1 in the Appendix shows the development of the average shares of non-western teachers and students and teachers and students from low-educated families for the period 2009-2019. These developments are shown separately for the four largest cities (G4) and the rest of the country (non G4). The key takeaway from these figures is that the shares are fairly constant over time with the exception of the share of students from low-income families which has been declining over time, especially in the G4.

Table 2: Means and standard deviations of schools' share of teachers and students with non-western migration background and from low-educated family for primary and secondary schools by city; 2019-2020

	Primary schools				**Secondary schools**			
	Teachers		Students		Teachers		Students	
A: Non-western	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Amsterdam	0.183	0.230	0.557	0.270	0.225	0.136	0.522	0.271
Rotterdam	0.144	0.170	0.580	0.242	0.202	0.163	0.464	0.272
The Hague	0.136	0.192	0.497	0.282	0.144	0.098	0.424	0.243
Utrecht	0.073	0.133	0.371	0.255	0.117	0.093	0.373	0.246
Rest of the country	0.022	0.066	0.184	0.171	0.050	0.046	0.162	0.133
B: Low-educated	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Amsterdam	0.101	0.089	0.204	0.176	0.117	0.050	0.246	0.246
Rotterdam	0.110	0.088	0.232	0.121	0.136	0.071	0.244	0.244
The Hague	0.104	0.086	0.223	0.129	0.100	0.040	0.232	0.232
Utrecht	0.095	0.065	0.138	0.188	0.080	0.058	0.184	0.184
Rest of the country	0.101	0.090	0.124	0.060	0.094	0.044	0.148	0.079
C: Corr(Nw;Le)								
Amsterdam	0.457		0.906		0.449		0.920	
Rotterdam	0.480		0.878		0.296		0.949	
The Hague	0.318		0.878		0.449		0.930	
Utrecht	0.433		0.909		0.620		0.920	

Note: Teachers are measured in full-time equivalents. Panel C reports correlations between a school's share of teachers (students) with non-western migration background and its share of teachers (students) from a low-educated family by city and level fo education.

5 Results

We present results in three subsections. Subsection 5.1 presents results for the dissimilarity index D for evenness, Subsection 5.2 presents results for the interaction index I for exposure, and Subsection 5.3 presents results for the correlations between schools'

compositions of teachers and students.

5.1 Evenness

Table 3 reports dissimilarity indices by city and year (2009 and 2019) for teachers (panel A) and students (panel B). Each panel reports dissimilarity indices for non-western migration background and low-educated family background for primary schools and for secondary schools. For each index we report its observed (raw) value and the systematic value which adjusts the observed value by subtracting the noise component.

The first row of Table 3 shows that D for segregation of primary school teachers by non-western migration background is around 0.5 in the four largest cities of the Netherlands. It is somewhat lower in Rotterdam than in the other cities. The values in this row are fairly similar for 2009 and 2019. Only in Utrecht there appears to be an increase in teacher segregation. The second row reports the values that are adjusted for sampling variation. This reduces the raw indices by 0.2 to 0.3. Systematic segregation of primary school teachers is highest in Amsterdam and The Hague. The adjusted values also make clear that the increase in the index in Utrecht cannot be attributed to a change of the noise component.[9,10]

Segregation by migration background is lower for secondary school teachers than for primary school teachers. This is true for both the raw and the adjusted values of D. The only exception is segregation in 2019 in Rotterdam. In Rotterdam the index for secondary school teachers increased considerably between 2009 and 2019 resulting in a value of the adjusted index that exceeds that for primary school teachers in Rotterdam in 2019.

Segregation of teachers by the education level of their siblings appears to be quite low. The raw values of D are already in the range that is considered low (less than 0.3) and the adjusted values are (with one exception) close to zero.

Panel B of Table 3 reports values for D for the segregation of students. The raw and adjusted values are not very different reflecting that the numbers of students per school

[9]Figure C1 in the appendix shows the year-to-year development of the different dissimilarity indices for segregation of teachers, and Figure D1 for segregation of students.

[10]For public school teachers in India, Fagernäs and Pelkonen (2017) report segregation levels of 0.612 by gender and 0.663 by caste. Because the average number of teachers per school is only 2.87, these high values are mostly attributable to noise.

are not very small such that the impact of sampling variation is minor. The values of D are (almost) all in the 0.3 to 0.6 range thereby pointing to moderate student segregation. This is true in all four cities, for both years, for both levels of education and for both background characteristics (migration and parents' education).

Table 3: Dissimilarity indices by city and year

	Amsterdam		Rotterdam		The Hague		Utrecht	
A: Teachers	2009	2019	2009	2019	2009	2019	2009	2019
Non-western, primary, raw	0.5212	0.5324	0.4478	0.4138	0.4924	0.5092	0.4977	0.5654
Non-western, primary, adj.	0.3145	0.3138	0.2102	0.1711	0.2643	0.2808	0.1837	0.2609
Non-western, secondary, raw	0.2964	0.2732	0.2669	0.3895	0.3074	0.3074	0.2923	0.3143
Non-western, secondary, adj.	0.1769	0.1865	0.1648	0.3028	0.1932	0.2191	0.1271	0.1943
Low-educated, primary, raw	0.2545	0.3163	0.2060	0.2760	0.2293	0.3111	0.2386	0.2616
Low-educated, primary, adj.	0.0627	0.0560	0.0073	0.0195	0.0363	0.0672	0.0259	0.0037
Low-educated, secondary, raw	0.1686	0.1959	0.0938	0.1816	0.0625	0.1806	0.1340	0.2532
Low-educated, secondary, adj.	0.0655	0.0911	0.0124	0.0892	0.0000	0.0662	0.0277	0.1198
B: Students								
Non-western, primary, raw	0.6021	0.5711	0.5857	0.5375	0.6454	0.6156	0.5417	0.4620
Non-western, primary, adj.	0.5442	0.5128	0.5298	0.4816	0.5962	0.5644	0.4898	0.4141
Non-western, secondary, raw	0.5338	0.5364	0.6246	0.5396	0.4806	0.4898	0.5018	0.4474
Non-western, secondary, adj.	0.4938	0.4986	0.5852	0.5012	0.4416	0.4536	0.4650	0.4170
Low-educated, primary, raw	0.5381	0.5159	0.5450	0.5098	0.5780	0.6156	0.4932	0.5192
Low-educated, primary, adj.	0.4741	0.4545	0.4743	0.4476	0.5144	0.5959	0.4285	0.4549
Low-educated, secondary, raw	0.4787	0.5364	0.5094	0.5268	0.4535	0.4898	0.5018	0.4474
Low-educated, secondary, adj.	0.4253	0.4840	0.4497	0.4814	0.4416	0.4536	0.4650	0.4170

Note: The table reports dissimilarity indices for teacher (panel A) and student (panel B) migration background and education background by level of education, city and year. Raw and adjusted indices are obtained using the procedures described in Subsection 3.1

5.2 Exposure

Table 4 reports interaction indices by city and year (2009 and 2019) for teachers (panel A) and students (panel B). Each panel reports interaction indices for non-western migration background and low-educated family background for primary schools as well as for secondary schools. For each index we report its observed (raw) value and the expected value if teachers/students are randomly assigned to schools. These expected values are based on the average of 100 simulations.

Teachers with a non-western migration background in primary schools in Amsterdam were in 2009 on average in schools where 46% of their colleagues did not have a migration background. If primary school teachers in Amsterdam would have been randomly assigned to schools, this percentage would have been 64%. The ratio of these two percentages equals 0.72, indicating that segregation causes non-western teachers in Amsterdam on average to be exposed to 28% fewer non-migrant teachers than would have been the case without systematic sorting. In 2019 this percentage equals 31% for primary school teachers in Amsterdam. These percentages are a bit lower for primary school teachers in the other cities, with a low 17% in 2009 in Rotterdam. The percentages are substantially lower for secondary school teachers, where all but one are below 10% (Rotterdam 2019 is the exception).

For interaction of teachers from families with different education background (measured as having at least one low-educated sibling), the exposure in primary schools is between 10% and 20% lower than would be the case under random assignment of teachers to schools. These percentages are highest in The Hague and lowest in Utrecht. Here again, the picture is remarkably different for secondary school teachers where the observed indices are very close to the indices that would occur under random assignment.

Panel B reports interaction indices for students. The observed (raw) indices for students are substantially below the indices that would occur under random assignment. The differences are largest for the indices for migration background in primary schools and smallest for the indices for education background in secondary schools.[11] All in all, we conclude that the picture that emerges from the interaction indices is not very different from that based on the dissimilarity indices.

[11] Figure C2 in the appendix shows the year-to-year development of the different interaction indices for segregation of teachers, and Figure D2 for segregation of students.

Table 4: Interaction indices by city and year

	Amsterdam		Rotterdam		The Hague		Utrecht	
A: Teachers	2009	2019	2009	2019	2009	2019	2009	2019
Non-western, primary, raw	0.4600	0.4345	0.5938	0.5484	0.5271	0.4894	0.6205	0.5689
Non-western, primary, rand.	0.6394	0.6325	0.7172	0.6948	0.6966	0.6639	0.7532	0.7534
Non-western, secondary, raw	0.5768	0.5187	0.6469	0.5253	0.6576	0.6186	0.7318	0.6521
Non-western, secondary, rand.	0.6289	0.5693	0.6815	0.6467	0.6831	0.6481	0.7443	0.6879
Low-educated, primary, raw	0.3832	0.5063	0.3926	0.4772	0.4004	0.4871	0.4562	0.5688
Low-educated, primary, rand.	0.4749	0.6017	0.4509	0.5460	0.4759	0.6155	0.5111	0.6307
Low-educated, secondary, raw	0.3982	0.4945	0.4031	0.5126	0.4014	0.4945	0.4618	0.5796
Low-educated, secondary, rand.	0.4156	0.5135	0.4236	0.5912	0.4097	0.5030	0.4643	0.6006
B: Students								
Non-western, primary, raw	0.1424	0.1409	0.1675	0.1670	0.1715	0.1509	0.2800	0.3108
Non-western, primary, rand.	0.2397	0.2205	0.2754	0.2470	0.3210	0.2573	0.4657	0.4566
Non-western, secondary, raw	0.1990	0.1817	0.2129	0.2262	0.2701	0.2311	0.3227	0.3226
Non-western, secondary, rand.	0.3049	0.2716	0.4108	0.3785	0.3948	0.3272	0.4836	0.4532
Low-educated, primary, raw	0.1938	0.3221	0.1302	0.2429	0.1552	0.2345	0.2567	0.3954
Low-educated, primary, rand.	0.2440	0.3601	0.2006	0.2971	0.2398	0.2998	0.3642	0.4831
Low-educated, secondary, raw	0.1421	0.3017	0.1086	0.2591	0.1193	0.2701	0.1817	0.3620
Low-educated, secondary, rand.	0.1994	0.3470	0.2607	0.3783	0.1897	0.3317	0.2801	0.4431

Note: The table reports interaction indices for teacher (panel A) and student (panel B) migration background and education background by level of education, city and year. Raw and random indices are obtained using the procedures described in Subsection 3.2

5.3 Correlation between teacher and student composition

In the previous subsections we have established that both teachers and students in each of the four largest cities in the Netherlands are segregated across schools. This is true for segregation by migration background as well as for segregation by education background. Teachers segregation by education background is, however, not much larger than what could be expected under random assignment of teachers to schools. In this subsection we examine to which extent the composition of teachers in schools in correlated with the composition of students in schools. To this end, we compute Pearson correlation coefficients of the share of teachers from a certain group in a school and the share of students from that group in the same school.[12]

Results are reported in Table 5. The shares of teachers and students with a migration background in a school are clearly very highly correlated. In primary schools the correlation coefficients are between 0.53 in Rotterdam and 0.66 in Amsterdam, and in secondary schools they are all close to 0.8.

For the correlations of the shares of teachers and students from low-educated families, the picture is more diverse. In Amsterdam, The Hague and Utrecht, these these correlations are of moderate size (around 0.3) for primary schools and are quite substantial (0.48 to 0.67) for secondary schools. The correlations for schools in Rotterdam contrast with this, being equal to 0.16 for primary schools and 0.15 for secondary schools. This concurs with the results in Tables 3 and 4 where Rotterdam is the only city where the observed values of D and I for teacher segregation by education background are almost indistinguishable from the values under random assignment. It is also consistent with the lower correlation between migrant and education background of teachers in secondary schools in Rotterdam than elsewhere (cf. panel C of Table 2).

[12] To be precise, the expression of the correlation coefficient is: $r_{tp} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{t_i-\bar{t}}{s_t}\right)\left(\frac{p_i-\bar{p}}{s_p}\right)$, where n is the number of schools, t_i is the observed share of migrant teachers in school i, $\bar{t}$ is the average of the observed shares of migrant teachers in schools in the city, p_i is the observed share of migrant pupils in school i, $\bar{p}$ is the average of the observed shares of migrant pupils in schools in the city, and s_t and s_p are the sample standard deviations of the observed shares of migrant teachers and pupils, respectively.

Table 5: Correlations between teacher and student school population 2009-2020

Primary schools	Amsterdam	Rotterdam	The Hague	Utrecht
% Migrant NW	0.6596	0.5276	0.5840	0.6227
% Low-educated	0.3398	0.1610	0.2444	0.3117
Secondary schools	Amsterdam	Rotterdam	The Hague	Utrecht
% Migrant NW	0.7963	0.8012	0.7773	0.7941
% Low-educated	0.5135	0.1456	0.4754	0.6722

Note: Correlations conditional on the city and school year, all significant at the 0.01 level.

A possible reason for the high correlations between schools' shares of teachers and students with a migration background, is that teachers and students from the same groups live in the same neighborhoods and study and work in the schools closest to where they live. Table A1 in the appendix shows that this is not the case. The table reports mean home-school distances in primary and secondary education by city separately for teachers and students. While primary school students live on average one kilometer away from the school they attend, this is four to eight kilometers for their teachers. The mean distances in secondary education are three to four kilometers for students and eight to 16 kilometers for teachers. The differences in distances between teachers and students indicate that students and teachers in the same school do not live in the same neighborhoods.[13]

A reason for the high correlations between teacher and students background in secondary schools is related to ability tracking. The vocational tracks attract larger shares of students with a migrant background and from low educated families than the college and university tracks. At the same time the educational requirements to be certified as a teacher in the vocational tracks differs from those to teach in the college and university tracks. It may therefore be that the correlations between teacher and student characteristics in secondary schools can be attributed to between track differences. Table A4

[13]Table A1 also reports mean distances for teachers to the nearest schools. These distances are around 0.3 km for primary school teachers and around 1 km for secondary school teachers. Table A2 reports dissimilarity indices for the counterfactual that all teachers are employed by the school nearest to their home address, and Table A3 reports what correlations between schools' teacher and student shares would be in that case.

shows that dissimilarity indices for students remain substantial even within tracks.[14] Table A5 shows correlations between teacher and students shares at the school×track level. These correlations remain substantial for migration background but are lower than the overall correlations in the third row of Table 5. For education background, the correlations vary substantially with a low -0.19 for the vocational-elementary track in Rotterdam and a high 0.71 for the same track in Utrecht.

6 Conclusion

Using register data from teachers in the four largest cities in the Netherlands, we have documented that teachers are segregated across schools by their migration background and their social background as measured by the education level of their siblings. Segregation by migration background is sizable even after adjusting for segregation that can be attributed to chance. Segregation by education background appears to be not much larger than what could be expected under random assignment of teachers to schools.

When we compute correlations between schools' shares of teachers and students from the same group, it turns out that these correlations are quite substantial in most cases. As we mentioned in the Introduction, it is unclear whether these high correlations are desirable or not. High correlations between teacher and student background are desirable to the extent that students benefit from being taught by "teachers like them". The high correlations may, however, at the same time frustrate policies aimed at desegregation of students. As documented by Oosterbeek et al. (2021), school segregation of students (in Amsterdam) is for a large part due to students from different backgrounds preferring different schools. Schools with a more diverse teacher workforce may be in a better position to attract a diverse student body.

References

Allen, R., Burgess, S., Davidson, R., and Windmeijer, F. (2015). More reliable inference for the dissimilarity index of segregation. *Econometrics Journal*, 18(1):40–66.

[14] We cannot calculate dissimilarity indices for teachers within tracks because the register data do not contain information about the track in which someone teaches

Billings, S. B., Deming, D. J., and Rockoff, J. (2013). School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg. *Quarterly Journal of Economics*, 129(1):435–476.

Egalite, A. J., Kisida, B., and Winters, M. A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review*, 45:44–52.

Fagernäs, S. and Pelkonen, P. (2017). Where's the teacher? How teacher workplace segregation impedes teacher allocation in India. IZA DP, No. 10595.

Frankel, D. M. and Volij, O. (2011). Measuring school segregation. *Journal of Economic Theory*, 146(1):1–38.

Frankenberg, E. (2008). The segregation of american teachers. *Education Policy Analysis Archives*, 17(1):1–45.

Gershenson, S., Hansen, M., and Lindsay, C. A. (2021). *Teacher Diversity and Student Success: Why Racial Representation Matters in the Classroom.* ERIC.

Gershenson, S., Holt, S. B., and Papageorge, N. W. (2016). Who believes in me? the effect of student–teacher demographic match on teacher expectations. *Economics of Education Review*, 52:209–224.

Goldhaber, D., Theobald, R., and Tien, C. (2019). Why we need a diverse teacher workforce. *Phi Delta Kappan*, 100(5):25–30.

Grissom, J. A., Rodriguez, L. A., and Kern, E. C. (2017). Teacher and principal diversity and the representation of students of color in gifted programs: Evidence from national data. *Elementary School Journal*, 117(3):396–422.

Massey, D. S. and Denton, N. A. (1988). The dimensions of residential segregation. *Social Forces*, 67(2):281–315.

Massey, D. S. and Tannen, J. (2015). A research note on trends in black hypersegregation. *Demography*, 52(3):1025–1034.

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

Prokic-Breuer, T. and Vermeulen, S. (2020). Teacher-student assortative matching and student achievement.

Rao, G. (2019). Familiarity does not breed contempt: Generosity, discrimination, and diversity in Delhi schools. *American Economic Review*, 109(3):774–809.

Reardon, S. F. and Owens, A. (2014). 60 years after brown: Trends and consequences of school segregation. *Annual Review of Sociology*, 40(1):199–218.

Villegas, A. M. and Irvine, J. J. (2010). Diversifying the teaching force: An examination of major arguments. *Urban Review*, 42(3):175–192.

A Additional tables

Table A1: Mean travel distance in km to school-location by city and year

	Amsterdam		Rotterdam		The Hague		Utrecht	
A: Teachers	2009	2019	2009	2019	2009	2019	2009	2019
Primary schools - observed	5.3	4.9	6	5.3	4.8	4.1	7.6	6.4
Primary schools - nearest	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
Share observed nearest	4.2%	3.7%	4.1%	4.1%	3.9%	3.7%	4.2%	3.6%
Secondary schools - observed	12.5	11.6	10	10	8	8.2	15.8	12.3
Secondary schools - nearest	1.1	1.0	1.1	1.2	1.0	1.0	0.9	1.0
Share observed nearest	4.9%	5.0%	5.6%	4.3%	5.6%	4.3%	3.7%	4.5%
B: Students								
Primary schools	1.2	1.3	1.0	1.2	1.1	1.2	1.0	1.0
Secondary schools	3.6	3.4	3.3	3	3.2	3.1	4	3.8

Note: For teachers we only considered the distance to schools (BRIN4) with one location (BRIN6), because the teacher data is collected at the school level (BRIN4). The mean teacher distance is weighted by the job size (FTE). The counterfactual analysis assumes that each teacher is employed in the school nearest to their home address.

Table A2: Counterfactual Dissimilarity indices teachers

	Amsterdam		Rotterdam		The Hague		Utrecht	
Primary schools	2009	2019	2009	2019	2009	2019	2009	2019
Nearest* non-western	0.46	0.47	0.50	0.46	0.54	0.57	0.41	0.41
Nearest low-educated	0.25	0.27	0.27	0.34	0.27	0.35	0.24	0.36
Zipcode** non-western	0.46	0.44	0.39	0.37	0.49	0.53	0.39	0.36
Zipcode low-educated	0.19	0.26	0.16	0.25	0.23	0.25	0.20	0.25
Secondary schools	2009	2019	2009	2019	2009	2019	2009	2019
Nearest non-western	0.36	0.38	0.40	0.40	0.27	0.28	0.28	0.31
Nearest low-educated	0.24	0.22	0.10	0.24	0.09	0.11	0.17	0.21
Zipcode non-western	0.41	0.43	0.34	0.37	0.43	0.42	0.43	0.38
Zipcode low-educated	0.23	0.28	0.17	0.25	0.16	0.29	0.15	0.22

*A counterfactual analysis where each teacher works at the nearest school. **A counterfactual analysis where there is only 1 school in each zipcode that all teachers attend, this captures residential segregation of teachers.

Table A3: Counterfactual correlations between teacher and student school population 2009-2020

Primary schools - nearest*	Amsterdam	Rotterdam	The Hague	Utrecht
% migrant NW	0.55	0.43	0.73	0.48
% low educated	0.28	0.19	0.41	0.33
Secondary schools - nearest	Amsterdam	Rotterdam	The Hague	Utrecht
% migrant NW	0.42	0.21	0.56	0.19
% low educated	0.34	0.13	0.26	0.01

*A counterfactual analysis where each teacher works at the nearest school.

Table A4: Dissimilarity indices students by track

	Amsterdam		Rotterdam		The Hague		Utrecht	
Vocational elementary	2010	2019	2010	2019	2010	2019	2010	2019
Migrant NW	0.57	0.59	0.69	0.62	0.58	0.55	0.66	0.56
Low-educated	0.45	0.34	0.57	0.40	0.15	0.08	0.72	0.55
Vocational theoretical	2010	2019	2010	2019	2010	2019	2010	2019
Migrant NW	0.56	0.47	0.59	0.54	0.45	0.47	0.35	0.46
Low-educated	0.41	0.39	0.51	0.44	0.28	0.24	0.36	0.38
College	2010	2019	2010	2019	2010	2019	2010	2019
Migrant NW	0.40	0.45	0.57	0.47	0.39	0.42	0.31	0.29
Low-educated	0.36	0.39	0.57	0.38	0.35	0.31	0.19	0.22
University	2010	2019	2010	2019	2010	2019	2010	2019
Migrant NW	0.28	0.31	0.40	0.37	0.26	0.37	0.20	0.17
Low-educated	0.24	0.32	0.38	0.35	0.35	0.31	0.06	0.05

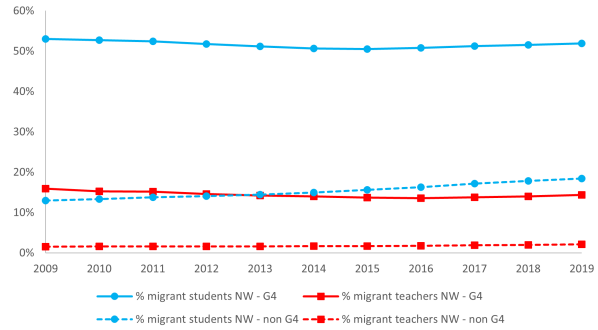
*School advice is collected in t-1 so observation period starts in 2010. Students with other tracks than the above or no track information aren't considered in this analyses.

Table A5: Correlations between teacher and student school population 2010-2020 by track

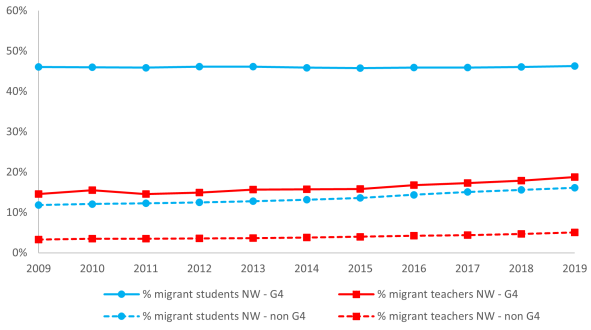
Vocational elementary	Amsterdam	Rotterdam	The Hague	Utrecht
% migrant NW	0.71	0.82	0.84	0.78
% low educated	0.31	-0.19	0.21	0.71
Vocational theoretical	Amsterdam	Rotterdam	The Hague	Utrecht
% migrant NW	0.67	0.55	0.76	0.43
% low educated	0.49	0.09	0.02	0.38
College	Amsterdam	Rotterdam	The Hague	Utrecht
% migrant NW	0.72	0.62	0.57	0.35
% low educated	0.51	-0.14	0.23	0.17
University	Amsterdam	Rotterdam	The Hague	Utrecht
% migrant NW	0.71	0.47	0.41	0.49
% low educated	0.59	-0.01	0.23	0.29

*School advice is collected in t-1 so observation period starts in 2010. Students with other tracks than the above or no track information are not considered in this analyses.

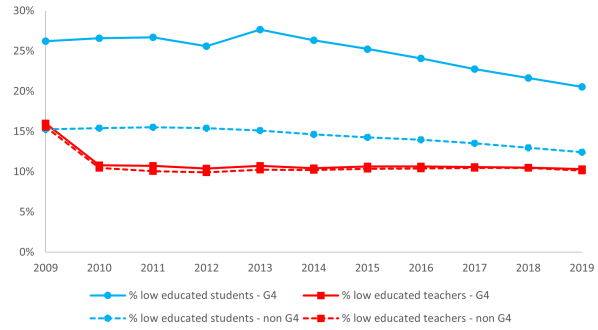
B Trends teacher and student population



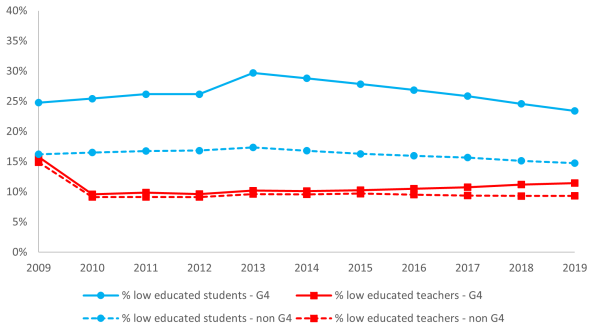
(a) non-western primary school



(b) non-western secondary school



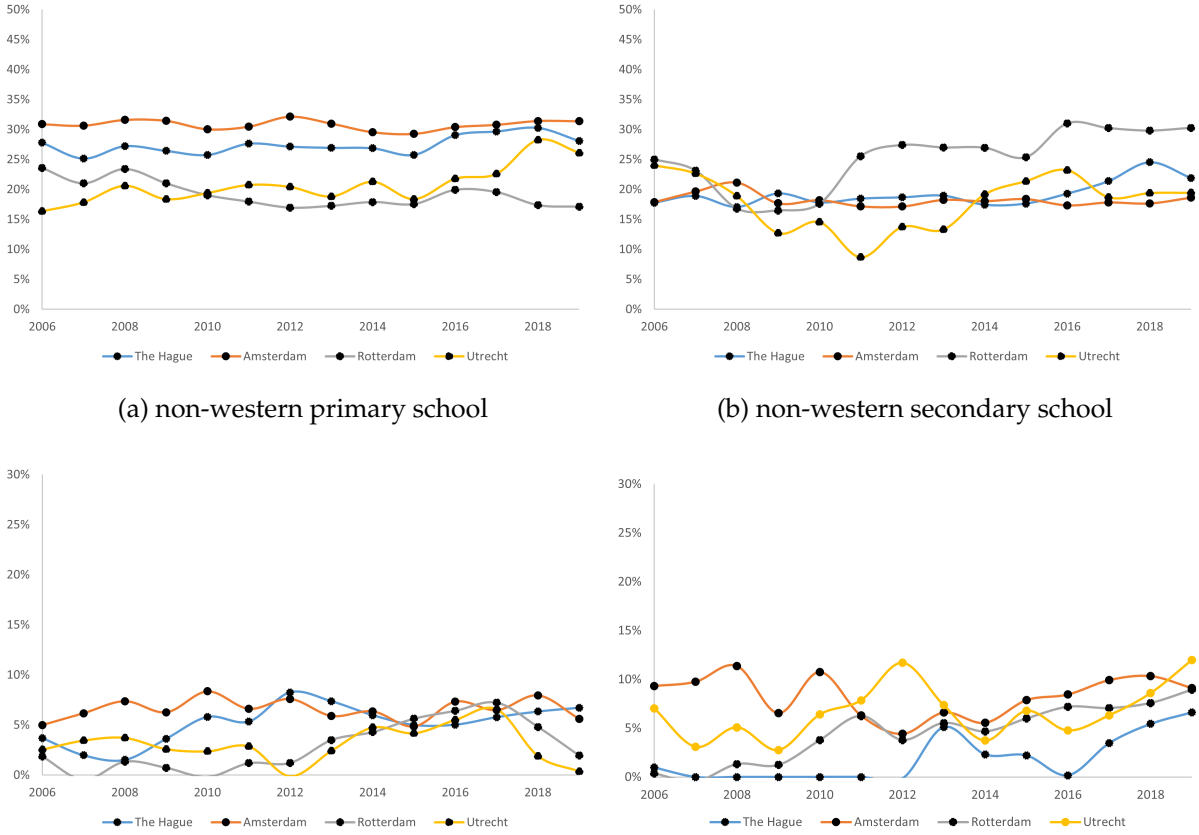
(c) low-educated primary school



(d) low-educated secondary school

Figure B1: Trends in teacher and student population

C Teacher segregation indices



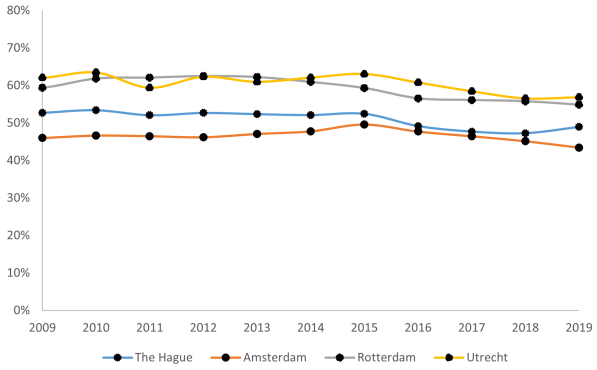
(a) non-western primary school

(b) non-western secondary school

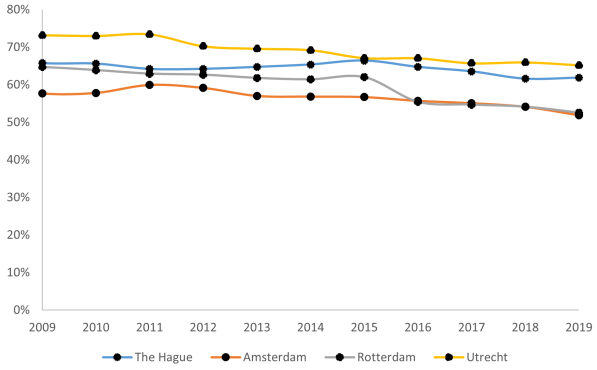
(c) low-educated primary school

(d) low-educated secondary school

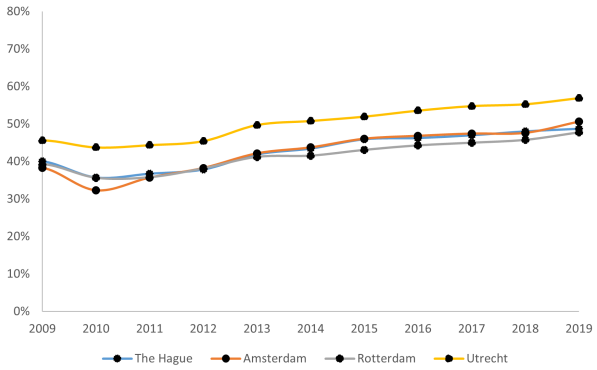
Figure C1: Adjusted Dissimilarity Index for teachers



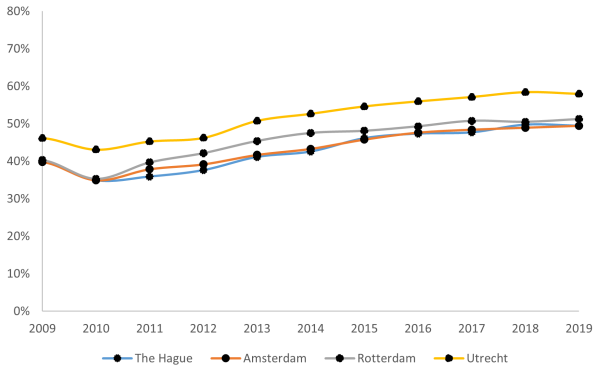
(a) non-western primary school



(b) non-western secondary school



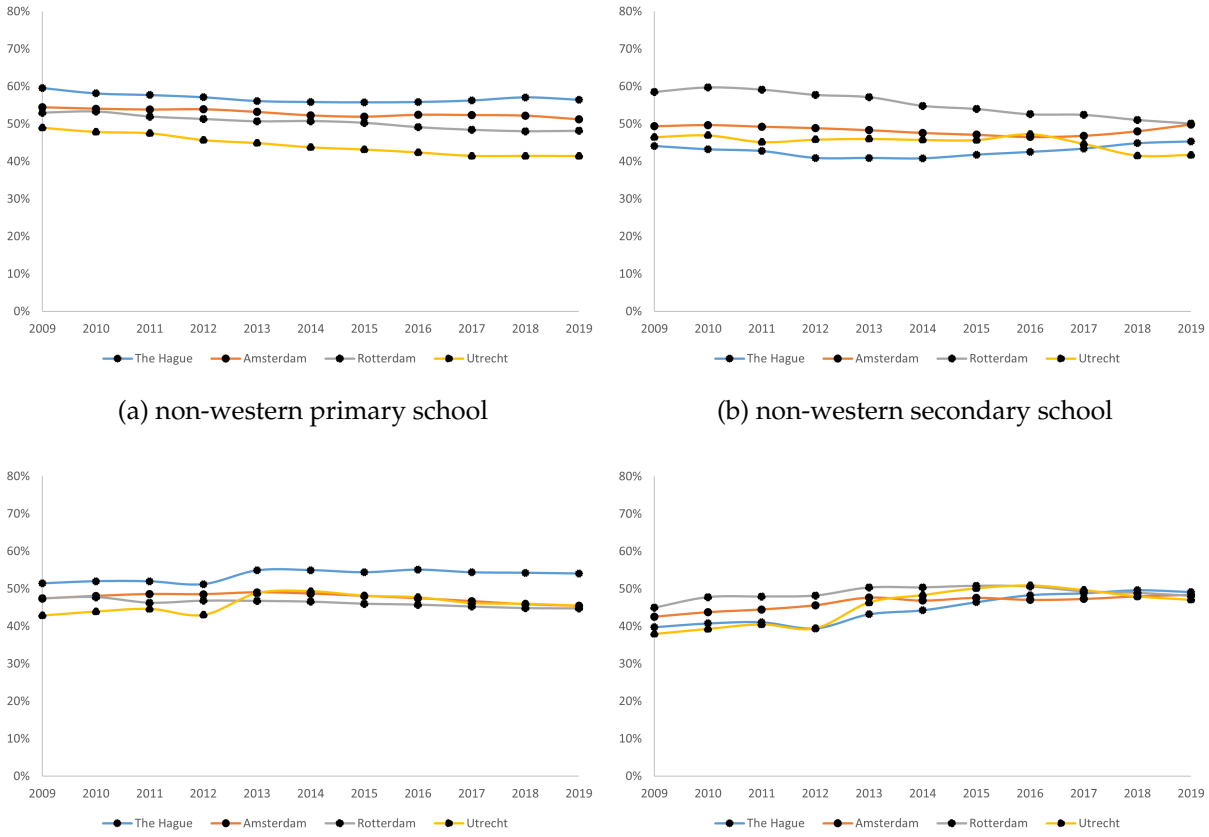
(c) low-educated primary school



(d) low-educated secondary school

Figure C2: Interaction Index for teachers

D Student segregation indices



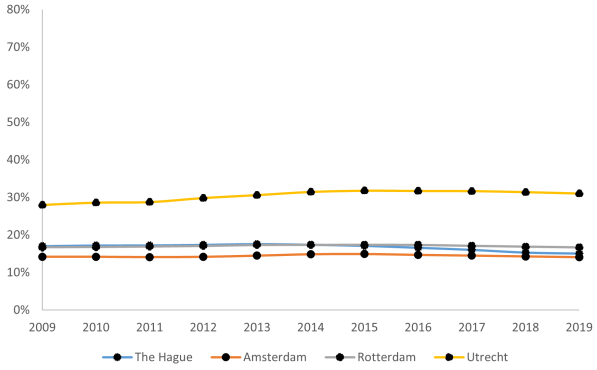
(a) non-western primary school

(b) non-western secondary school

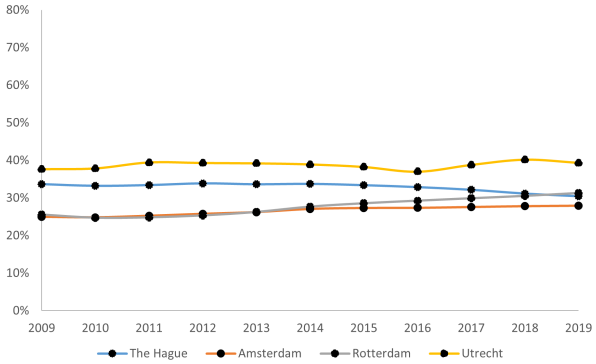
(c) low-educated primary school

(d) low-educated secondary school

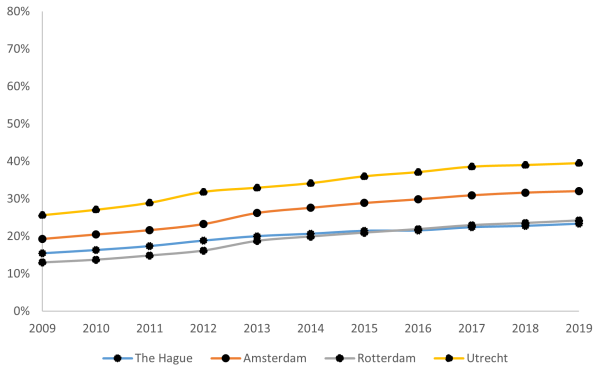
Figure D1: Adjusted Dissimilarity Index for students



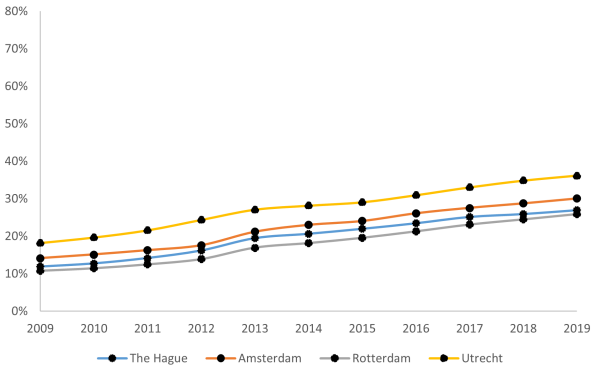
(a) non-western primary school



(b) non-western secondary school



(c) low-educated primary school



(d) low-educated secondary school

Figure D2: Interaction Index for students