Educational Research and Evaluation, 2013
http://dx.doi.org/10.1080/13803611.2013.788851

Routledge
Taylor & Francis Group

Ethnic composition of the class and educational performance in primary education in The Netherlands

Gert-Jan M. Veerman[a]*, Herman G. van de Werfhorst[a] and Jaap Dronkers[b]

[a]*Department of Sociology and Anthropology, University of Amsterdam, Amsterdam, The Netherlands;* [b]*Research Centre for Education and the Labour Market, Maastricht University, Maastricht, The Netherlands*

(*Received 20 April 2012; final version received 5 February 2013*)

This article examines the effect of the ethnic composition in the school class on school performance in primary education, using COOL 2008 data for The Netherlands. We make an important distinction between the *proportion* of migrant children and the *diversity* with regard to the different ethnic groups in a school class. Due to the strong correlation between these 2 variables, we employ a residualized score of diversity on the proportion of migrants. The diversity indicator, which indicates the level of diversity given a particular share of migrant children, is negatively related to reading comprehension in Grade 8. For other grade years, we find little support for negative effects of diversity net of the share of migrants in a class.

Keywords: academic achievement; ethnic composition; ethnic diversity; multicollinearity

Introduction

The relationship between the ethnic composition of schools and pupils' achievement is of growing interest to European researchers (Agirdag, Van Houtte, & Van Avermeat, 2012). Recently, the studies of Dronkers and Van der Velden (in press) and Maestri (2011b) made the interesting distinction between the ethnic *share* and ethnic *diversity* in a school or grade. The ethnic share refers to the proportion of migrant children in a class (independent of ethnic group), whereas ethnic diversity refers to the composition in the class in terms of the number and size of different ethnic groups. This distinction is conceptually relevant, but the current literature is often blurred. Studies on the influence of "diversity" have in fact focused mostly on the proportion of migrant children, rather than on the diversity within a class as such.

A few earlier studies focused on the distinction between the share of migrant children and the diversity among them. Maestri's research (2011b) found that ethnic diversity increases language and math test scores of pupils with an immigrant background in primary education in The Netherlands. Nevertheless, Dronkers and Van der Velden (in press) found, with data from a selection of Organisation for Economic Co-operation and Development (OECD) countries of secondary school pupils, that a greater ethnic diversity

*Corresponding author. Email: G.J.M.Veerman2@uva.nl

of schools has a significant negative effect on the learning performance of migrants.[1] Meastri (2011b) explained in her study how differences in research designs probably influence the results.

Despite the growing awareness of the conceptual distinction between the share of migrants and the diversity in terms of the different groups, it seems that researchers interested in diversity have also often left out the share of migrants from their empirical models (Demanet, Agirdag, & Van Houtte, 2011; Herweijer, 2011). Related literature on the impact of ethnic diversity in geographical units on social cohesion has completely ignored the share of migrants as a variable of interest (Gijsberts, Van der Meer, & Dagevos, 2012; Lancee & Dronkers 2011; Tolsma, Van der Meer, & Gesthuizen, 2009). This is unfortunate, as the conceptual distinction between proportions and diversity warrants closer inspection of their isolated effects on individual outcomes such as student learning. An important reason why researchers have often examined only one of the two concepts is presumably that these are strongly correlated, leading to problems of multicollinearity.

Therefore, we are interested in the partial effects of ethnic diversity and the proportion of migrant pupils on school performance in mathematics and reading comprehension. We study this using rich survey data gathered among pupils in different grades in primary education in The Netherlands. We aim to contribute to the literature in three ways. First, we will explicitly demonstrate the methodological problem of including both the ethnic share and ethnic diversity in one model, and offer a solution to deal with this problem. This is necessary, because ethnic diversity and ethnic share are strongly correlated. Second, we measure ethnic diversity at the school class level instead of the school or grade level. We prefer class level, because most theories about the influence of ethnic diversity on school performances refer to the class level. Finally, we distinguish both the proportion of migrants and ethnic diversity in our model, using two distinct conceptualizations of diversity (Herfindahl's diversity index and the number of ethnic groups in the class). These two variables isolate different processes concerning diversity, and should be studied likewise.

The Dutch situation

The Dutch case is very interesting for the research question for several reasons. First, in The Netherlands there is a rich dataset (*Cohort Onderzoek Onderwijsloopbanen*, COOL) of 38,060 pupils from three different grade levels in primary education. Therefore, the group sizes of the different ethnic groups are large enough to distinguish 11 different ethnic groups at three different grades in primary school. Second, the database contains 550 schools with a wide range of ethnic compositions, containing both schools with a high proportion of migrants and a low proportion of migrants varying with percentages of immigrants between 0 and 100% and schools in both rural and urban settings. Third, the percentage of foreign-born citizens in The Netherlands (11%) is comparable with the percentages of Norway (11%) and The United Kingdom (11%) and just below the percentage of France (12%) and Germany (13%) (OECD, 2011). Therefore, analyses of the Dutch case give indications of the effects of the ethnic composition that could occur in other European countries. Nevertheless, we should take into account the specific Dutch educational policy and the migration history.

Dutch educational policy

Earlier (comparative) studies characterized Dutch education policy by three essential elements (Dronkers, 2004; Ladd & Fiske, 2009). First, families have freedom of school

choice, which makes it possible to avoid schools with high shares of migrant pupils without the need to move house (Van Houtte & Stevens, 2009). This may explain the high level of ethnic segregation in Dutch primary schools, as has been documented by Ladd, Fiske, and Ruijs (2010) and Karsten (2010). Second, by constitution parents have the right to establish their own schools under equal funding rights as public schools (and equal forms of government control concerning school quality). Originating from the "school struggle" in the early 20th century, the constitutional right enabled schools of religious denominations to establish schools under state funding. Third, the funding of schools from the central government is based on weighted pupil funding. Dutch schools with large concentrations of disadvantaged pupils receive more resources than schools with pupils of middle-class backgrounds. The funds per pupil depend on parents' education (and until 2006 also on ethnic background; Ladd & Fiske, 2009).

Migration history

The Netherlands has a long migration history. Starting with the post-war period, the first big migration wave came from the former Dutch colony of the East Indies (Lucassen & Penninx, 1997). The Turks and Moroccans started to migrate to The Netherlands during the 1960s after the influx of other "guest workers" from Italy, Greece, Portugal, and Spain. The migration from Suriname accelerated in the 1970s, when it became independent from The Netherlands. Later, also the influx of migrants from the Netherlands Antilles, in the Carribean, grew. More recent groups of refugees and those of "other" Western origin have been migrating to The Netherlands since the 1990s. These new, relatively small groups have become visible in the recent COOL dataset and have made the Dutch classrooms more ethnically diverse than previously.

Theory and hypotheses

Ethnic composition and school performances

Why does the ethnic composition of schools or classes affect student performance? Studies propose several explanations for the effect of the ethnic composition. Two distinct perspectives are relevant: the *teaching* and the *peer group perspective*.

First, it has been argued from a *teaching perspective* that a higher proportion of migrants can lead to lower educational performances due to lowering the standards (Rosenthal & Jacobsen, 1968). The proportion of minority pupils may, however, also influence the educational performances positively, because teachers are likely to specialize to the needs of the minority pupils (Peetsma, Van der Veen, Koopman, & Van Schooten, 2006). Although the growth of the proportion of migrants may lead to specialization, it may also lead to teaching problems concerning instructional time for a greater number of ethnic groups (Dronkers & Van der Velden, in press; Maestri, 2011a). Moreover, teachers need to adapt their teaching style to the needs of a diverse set of pupils (Van Ewijk & Sleegers, 2010b).

Second, from the *peer group perspective* researchers propose that ethnic diversity can enrich students through communication, for instance, if the information about the culture of one ethnic group is relevant for the other group (Lazear, 1998). Also, the size of the ethnic groups could influence school performance, as smaller ethnic groups have stronger incentives to adapt to the majority culture (Lazear, 1999). Smaller ethnic groups could then lead to better understanding instructions because the instructional language is mostly determined

by the majority (Maestri, 2011a). However, the existence of small ethnic groups may also lead to lower school achievement due to a mechanism of reduced feelings of ethnic identification (O'Reilly, Williams, & Barsade, 1997). Interethnic contacts may lead to more interethnic tensions, which could negatively influence academic performance (Hoxby, 2000). Finally, the pupils' language development may be inhibited by a higher number of interethnic contacts due to fewer contacts with pupils having the host country language as their mother tongue (Driessen, 2002).

The various explanations of the effect of the ethnic composition on school performances show that researchers refer to four distinct elements of the ethnic composition: the *proportion of migrants*, the *number of ethnic groups*[2], the *size of the ethnic groups*, and the *number of interethnic contacts.*

Proportion of migrants

The score of migrant students might be positively influenced by the specialization to the needs of migrant students of the teachers in school classes with a high ethnic share. Nevertheless, two other mechanisms might explain a negative effect of a higher ethnic share on school performance. First, besides this positive effect, we expect a negative influence of lower expectations (Rosenthal & Jacobsen, 1968) of teachers in classes with a higher share. Second, a higher proportion of migrants leads to a lower chance of contacts with native pupils with higher skills of the destination country language. Consequently, we expect in our *proportion of migrants hypothesis* for both migrant and native pupils in primary education due to the two negative mechanisms that: A higher proportion of migrants is associated to lower school performance.

Number of origin groups

A study of Driessen (2002) examined the proportion of ethnic minority youth and *the number of origin groups* as indicators of school composition. The study shows for both math and reading no influence of the number of origin groups on the school performances in Grades 4 and 8 in Dutch primary education. The number of origin groups does not directly relate to the group sizes and the opportunity structure of interethnic contacts. Rather, the number of groups is usually considered important because of the instructional problems that may arise from teaching a large number of different groups. The study of Driessen shows no significant effect of the number of origin groups on the school performances in primary education. One reason why no association was found between student achievement and the number of groups may be that teachers are able to specialize to cater the needs of the different groups of migrant pupils. Another reason could be that the instructional needs do not differ a lot between different origin groups. Furthermore, Lazear (1998) argued that ethnic groups can enrich students if the information about the culture of one ethnic group is relevant for the other group. Nevertheless, the information that other origin groups could supply is probably not relevant for all school performances. Furthermore, if the information is relevant for the other group, the use of this information is only structurally implemented in some curricula (Svalberg, 2007).

Yet, with our data we would like to put the *educational instruction hypothesis* to another test, which states: There is a negative association between the number of origin groups and school performance.

Origin diversity

Recent studies have also taken into account *the composition of the class with regard to the sizes and the number of different origin groups* using an *origin diversity* variable based on the Herfindahl index. Origin diversity measured this way is based on both the proportions of the separate origin groups and the number of the origin groups (Dronkers & Van der Velden, in press). A low origin diversity index refers to fewer relatively small origin groups and a high origin diversity index to more relatively large origin groups. Because origin diversity contains both the *number* and the *size* of the origin groups, origin diversity refers both to the earlier mentioned educational instruction mechanisms and to peer group influences. Larger origin groups might lead to both positive and negative peer group influences on school performances. Larger origin groups might lead to positive influences on school performances due to a mechanism of stronger feelings of ethnic identification (O'Reilly et al., 1997). Nevertheless, these larger ethnic groups also give smaller incentives to adapt the culture of the destination country. According to Maestri (2011b), this may lead to more problems in understanding instructions. As we demonstrate in Appendix 1, origin diversity also directly relates to the *relative number of possible interethnic contacts.* This relatively higher number of possible interethnic contact could lead to more interethnic tensions and conflicts (Hoxby, 2000).

We especially expect a decisive influence of the combination of both negative peer group and instruction mechanisms of diversity on student achievement. Therefore, we formulate the *diversity hypothesis* as follows: A higher origin diversity is associated to lower school performance for both migrant and native pupils.

Data and variables

The COOL data

The analyses have been carried out using the Cohort Research on Educational Careers (*Cohort Onderzoek Onderwijsloopbanen*, COOL) funded by The Netherlands Organisation for Scientific Research (NWO) and the Dutch Ministry of Education, Culture and the Sciences.[3] The information in the COOL survey data is both rich and relatively large. The primary school information has been gathered by two Dutch institutes: the Kohnstamm Institute (KI) and the Institute for Applied Social Sciences (ITS). For our study, we used a school-based survey of 38,060 pupils from 550 primary schools in The Netherlands (Driessen, Mulder, Ledoux, Roeleveld, & Van der Veen, 2009). We used the first available wave in primary education, which was collected in 2008 and includes results of pupils in Grades 2, 5, and 8 (ages mostly around 5/6, 8/9 and 11/12).[4]

We had the use of the information of 36,796 pupils; this means that we lost 2% of our respondents due to missing values on parental education and origin. Furthermore, for 8% of the remaining native pupils and 9% of the pupils with a migration background, the mathematics test score is missing. In Appendix 2, we show the difference between the pupils with math test scores and without math test scores on a number of background variables. With respect to the pupils without valid information on math scores, the first-generation migrant pupils in Grade 5 and migrant pupils with parents with a lower education in Grade 8 are overrepresented. Nevertheless, all origins and education levels are represented in both the missing group and the available group. Therefore, we expect hardly any influence of the missing data on our outcomes.

It should be noted that the data have been gathered by grade and not by school class. If pupils were part of "combination classes" of multiple grades, we collected group-level

information about the Grades 2, 5, or 8 fraction of the class. We have excluded all classes with fewer than five pupils to remove the cases with unreliable group-level variables, covering another percent of the pupils. Therefore, the total number of pupils with math scores in our database is 33,624.

Variables

Dependent variables

Academic performance. The dependent variable academic performance is a score on the math test developed by the national testing agency Cito. We expect stronger effects for the math scores to language scores, because in general math scores are more strongly related to school class influences than language scores (Creemers, 2007). Nevertheless, we have also included the analysis on reading comprehension, because Dronkers and Van der Velden (in press) mentioned that the results of language skills are more pronounced for pupils with an immigrant background.

The Cito math and reading comprehension tests are nationally standardized. The tests are taken twice every year in most primary schools in The Netherlands. The COOL dataset contains the first test of the school year. The test scores are used by teachers and researchers to monitor the development of the individual pupils. Furthermore, the Dutch Inspectorate of Education uses the test scores to assess and compare the quality of primary schools.[5]

Class level variables

Proportion of migrants of the school class. We computed the proportion of migrants of the school class using the percentage of migrant pupils in the class. This includes first- and second-generation migrants, using the official definition of Statistics Netherlands which is based on the country of birth of child and/or parent(s).

Origin diversity. Using the number of pupils per origin caught up in every class, we computed an inverted Herfindahl index of origin diversity. We calculated the index as follows: 1 – ((percentage origin group 1)^2 + (percentage origin group 2)^2 + … + (percentage origin group n)^2). Although we argued that the proportion of migrants and origin diversity are concepts that we should distinguish on theoretical and empirical grounds, the use of both variables could lead to a problem of collinearity. This problem occurs due to the strong correlation between proportion of migrants and the origin diversity of $r = 0.93$ at the school class level for natives and $r = 0.85$ for migrants. Therefore, a quadratic model was estimated on the class-level data, predicting diversity as a function of the proportion of migrants. We then took the residuals of this regression model, thereby measuring the difference between origin diversity as is observed in a class relative to the predicted diversity based on the quadratic model. The advantage of this method is that the residualized diversity measure is independent of the proportion of pupils with a migration background, as independence of the residual with X variables is an assumption of ordinary least squares regression. Our measurement thus does not assess diversity per se, but the level of diversity given a particular proportion of children of non-Dutch descent. Such a model may underestimate the impact of diversity, yet we do think that this measure associates directly to the conceptual distinction between the share of migrants and the diversity among them. Classes with similar numbers of migrant pupils vary with regard to the diversity of the groups, and

our measure conceptualizes such a "conditional" interpretation of diversity. The residualization process may also be done in reversed fashion (with proportion of migrants residualized on diversity), which would give more leeway to Diversity as a predictor variable. The resulting interpretation of such a reversed residualization would be that, given a particular level of diversity, a larger proportion of migrants may be related to school performance. Such an interpretation is, in our view, slightly less desirable.

In Appendix 3, we show how origin share and origin diversity are related and how we computed the residuals of origin diversity. Furthermore, we will show in the section ***A test on multicollinearity*** how both the use of origin share and non-residualized origin diversity in one model leads to possible problems of collinearity. This section will also show that the use of both origin share and *residualized* origin diversity solves the collinearity problem.

Number of origin groups per class. We calculated the number of origin groups per class, using information at the individual level about the origin of the pupils. Similar to our approach on the diversity index, we also calculated the number of origin groups residualized to the proportion of migrants in a class. This is necessary as also the number of migrant groups is strongly correlated to the share of migrants.

Percentage of parents with higher education in the class. The COOL dataset contains two measurements of the parental education: the parental education according to the school and the parental education as mentioned by the parent in the parental questionnaire. The information of the first measurement is estimated by the school and frequently originated from the administration of the school (Driessen, Mulder, & Roeleveld, 2012). Although the answers of the parental questionnaire might be more recent and have more specified levels, the parental dataset leads to 34% missing values for the parental educational level. Furthermore, the response of the parental questionnaire is skewed on social ethnical background (Driessen et al., 2009). Consequently, the choice for the parental education according to the parental questionnaire would lead to a selection of missing data and problems to calculate the percentage of parents with higher education at the class level. Therefore, we prefer the educational measurement of the school that contains only to 5% missing values of the parental education. Nevertheless, we underpin that our indicator of the parental educational level is an estimated variable in the cases where schools did not originate their data from their administration. We calculated the percentage of parents with higher education in the class, defined as tertiary level vocational college (HBO) or university. In some classes, no information on parents' education was collected. We omitted those classes from the analysis (321 pupils).

Individual level variables

Origin background. Using the method applied by Levels and Dronkers (2008), we took the country of birth of the child, the father, and the mother as indicators for origin. If two of these three indicators had the same country of birth but not the country of destination, we took that country as origin country. However, when there were not two of the same classifications available, the country of birth of the mother was taken to represent the origin country. The COOL dataset contains also two measurements for the origin country: according to the school and as mentioned in the parental questionnaire. The parental questionnaire is, for instance, missing 36% of the information about the country of origin of the mother and the school measurement only 4%. Therefore, we prefer also for the origin country the measurement according to the school, because otherwise we

would lose a large number of pupils. Because of the small group sizes, we put Moluccan pupils together with Surinamese and Antillean pupils as former colonies, and Polish, former Soviet, and former Yugoslavia as Eastern Europe. We did compute the origin diversity and number of origin countries on the basis of all individual countries. The country of origin is missing for 943 pupils.

Migration generation. Using information on the countries of birth of the pupils and their parents, we constructed a dichotomous variable. We define first-generation migrants as pupils who were born abroad and whose parents were also born outside The Netherlands. We define second-generation migrants as pupils who were born in The Netherlands and of whom at least one parent was born abroad.

Parental educational level. Parental educational level is measured as the highest level of either of both parents according to the school. The educational levels that are distinguished are primary education (low); lower secondary education; upper secondary education (vocational, general, or academic); and tertiary education(vocational college or university). We include the dummy "parental educational level missing", representing around 1% of the pupils.

Female. We use a dichotomous variable to classify gender. Boys are the reference group. Of 1% of the pupils, we have no information on gender. The mathematics test results of this group did not significantly differ from those of the children whose gender was known. This group has been assigned the value 0.5 on the gender dummy.

Descriptive statistics

Table 1 reports the means and standard deviations for migrant pupils (first and second generation together) and native pupils of all variables in all years. We see that in all years math scores are lower for migrant pupils than for native pupils. The difference in the dataset in test scores between migrant pupils and native pupils declines between Years 5 and 8. The table furthermore shows that migrant pupils have classes with a higher level of origin diversity and a higher share of migrants.

Models and results

Analytical design

Given the nested structure of the data, with individual pupils nested in classes, which are nested in grades, we used multilevel analysis with three levels.

We ran separate analyses for native pupils and pupils with a migration background. Our first model (1a) in Table 2 only contains the proportion of migrants and the origin diversity of the class as predictor variables. In Model 1b, we have replaced the variable origin diversity by the residualized version of that variable. We use both Models 1a and 1b because these models show the influence of the possible multicollinearity on the parameter estimates. In the following section, we will show why we prefer Model 1b to Model 1a. In Model 2, we have added all explanatory variables except the residuals of origin diversity and the residuals of the number of origin groups. With this model, we can test whether the addition of the residualized variables in Models 3 and 4 improve the model fit. Our third model also contains the residualized origin diversity index. Finally, our last model

Table 1. Means and standard deviations for migrant students and native students with math scores[a].

	Grade 2				Grade 5				Grade 8			
	Migrants		Natives		Migrants		Natives		Migrants		Natives	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
% migrant students of class	63.15	28.10	14.57	19.85	63.56	29.46	13.51	18.81	62.64	29.62	13.95	18.77
Origin diversity of class	0.60	0.18	0.20	0.23	0.59	0.20	0.20	0.22	0.59	0.19	0.20	0.22
Residualized origin diversity of class	0.00	0.14	–0.00	0.08	0.00	0.17	0.00	0.08	0.01	0.15	0.00	0.08
% parents with tertiary education of class	18.96	18.51	33.58	22.05	16.76	16.88	32.49	20.36	16.14	15.22	30.02	19.69
Number of origins in class	4.48	1.55	2.27	1.50	4.87	1.63	2.52	1.62	4.97	1.69	2.66	1.73
Residualized origins in class	0.29	1.43	0.12	0.88	0.21	1.58	0.16	0.98	0.22	1.65	0.16	1.09
Education missing	0.01	0.10	0.01	0.09	0.02	0.13	0.01	0.12	0.02	0.14	0.01	0.12
Low parental education	0.28	0.45	0.02	0.13	0.30	0.46	0.02	0.12	0.30	0.46	0.02	0.12
Lower secondary parental education	0.23	0.42	0.16	0.37	0.24	0.43	0.18	0.38	0.25	0.43	0.20	0.40
Upper secondary parental education	0.30	0.46	0.48	0.50	0.30	0.46	0.46	0.50	0.27	0.45	0.47	0.50
Tertiary parental education	0.18	0.39	0.34	0.47	0.15	0.36	0.33	0.47	0.15	0.36	0.30	0.46
Female	0.47	0.50	0.49	0.50	0.50	0.50	0.50	0.50	0.48	0.50	0.49	0.50
First-generation migrant	0.06	0.23			0.09	0.28			0.14	0.34		
Turkish origin	0.24	0.43			0.27	0.44			0.28	0.45		
Moroccan	0.24	0.43			0.25	0.43			0.21	0.41		
Western	0.07	0.25			0.06	0.23			0.06	0.23		
Eastern Europe	0.05	0.21			0.05	0.21			0.05	0.22		
Chinese	0.01	0.11			0.01	0.12			0.02	0.12		
Iraqi	0.03	0.16			0.02	0.13			0.03	0.16		
Afghan	0.02	0.13			0.01	0.11			0.02	0.15		
Somali	0.03	0.16			0.02	0.13			0.01	0.11		
Other country	0.20	0.40			0.18	0.39			0.17	0.37		
Former colony	0.12	0.32			0.13	0.34			0.15	0.36		
Math score	51.87	11.88	58.56	13.04	63.60	15.72	71.96	14.97	114.42	9.52	117.12	9.40
N students	3426		8803		3056		8290		2719		7330	
N classes	719		991		505		677		438		580	
N schools	369		494		383		495		362		477	

[a]Only reported for students with valid scores on the mathematics test. Source: COOL 2008, own computation.

Table 2. The effects of the origin composition on math score of native students and migrant students: diagnosing collinearity.

	Grade 2				Grade 5				Grade 8			
	Model 1a		Model 1b		Model 1a		Model 1b		Model 1a		Model 1b	
		VIF		VIF		VIF		VIF		VIF		VIF
Native												
Constant	59.2**		59.3**		73.0**		73.3**		118.0**		117.6**	
	(0.4)		(0.4)		(0.4)		(0.4)		(0.3)		(0.3)	
% migrant students of class	–7.8**	7.0	–5.9**	1.0	–13.3**	7.4	–9.4**	1.0	–0.5	7.2	–5.2**	1.0
	(2.9)		(1.0)		(3.8)		(1.2)		(2.6)		(0.9)	
Origin diversity of class	1.8	7.0			3.9	7.4			–4.7	7.2		
	(2.7)				(3.5)				(2.4)			
Residualized origin diversity of class			–1.6	1.0			2.2	1.0			–8.7	1.0
			(6.7)				(10.2)				(6.9)	
Log likelihood	69259.1		69259.5		67777.7		67777.7		53006.3		53008.4	
N students	8803				8290				7330			
N classes	991				677				580			
N schools	494				495				477			
Migrant												
Constant	53.7**		53.6**		68.8**		68.4**		116.2**		115.6**	
	(0.9)		(0.7)		(1.1)		(0.8)		(0.7)		(0.5)	
% migrant students of class	–2.2	3.5	–2.5*	1.0	–6.4**	3.5	–7.3**	1.0	–0.4	3.7	–1.9*	1.0
	(1.7)		(1.1)		(2.1)		(1.4)		(1.4)		(0.9)	
Origin diversity of class	–0.5	3.5			–1.6	3.5			–2.3	3.7		
	(2.4)				(2.9)				(2.0)			
Residualized origin diversity of class			–0.1	1.0			–3.4	1.0			–2.0	1.0
			(3.8)				(5.0)				(4.1)	
Log likelihood	26248.5		26248.6		25229.2		25229.2		19803.8		19804.9	
N students	3426				3056				2719			
N classes	719				505				438			
N schools	369				383				362			

Notes: Standard errors between brackets. Significance: $^{**}p < 0.01$; $^{*}p < 0.05$.

Source: COOL 2008, own computation.

contains all explanatory variables, but we have replaced the residualized origin diversity score by the similarly residualized number of origin groups.

A test on multicollinearity

Due to the strong correlation between origin share and origin diversity, a diagnosis of multicollinearity was performed by examining the variance inflation factor (VIF).[6] The VIF quantifies the degree to which estimated standard errors of regression coefficients are influenced by linear relationships among predictor variables (Kleinbaum, Kupper, Nizam, & Muller, 2008).

Table 2 reports the results from our multilevel regression for Models 1a and 1b and ordinary least squares (OLS) VIF statistics as an indication to assess multicollinearity. The statistics show for the model that contains both the proportion of migrants and the origin diversity a maximum VIF of 7.2 for the native pupils and a maximum VIF of 3.7 for the migrants. As we expected, the VIF's of Model 1b are all one, because the residualized score on origin diversity does not correlate with the proportion of migrants.[7]

A diagnosis of multicollinearity using VIF scores depends on the chosen threshold of an acceptable VIF. Earlier research used a VIF of 4, 5, or 10 as rule of the thumb to indicate serious multicollinearity (O'Brien, 2007). High values of VIF lead to inflated standard errors of regression coefficients, impacting the statistical significance of regression coefficients. We also compare the standard errors and regression coefficients between Models 1a and 1b. Table 2 shows that the standard errors of the proportion of migrants for the natives meanly inflate between Models 1b and 1a with 199% and for the migrants on average with 54% due to the high correlation between the independent variables. Furthermore, the table shows for natives in Grade 8 a parameter of –5.2 for the proportion of migrants using the residualized variable of diversity, whereas the model with the high VIF scores shows an effect of –0.5. All model comparisons show higher standard errors and in most cases lower parameter estimates for the proportion of migrants if the model was estimated with the unresidualized measure of origin diversity. Consequently, we choose Model 1b with the residualized origin diversity index. Furthermore, in our fourth model we also take the residualized version of the number of origin groups, because origin share and the number of origin groups also correlate strongly.

Results

Proportion of migrants

Tables 3a and 4a show, for natives in all subjects, a negative effect of the proportion of migrants on school performance. For migrant pupils, the results of Tables 3b and 4b show varied outcomes in this regard. For math, Table 3b shows a significant effect of –3.7 or –3.8 only in Grade 5. Furthermore, the results on reading comprehension show for migrants no significant effect of the proportion of migrants on the test scores. Consequently, we reject the *proportion of migrants hypothesis* for the migrants except for math in Grade 5 and confirm the proportion of migrants hypothesis for the native pupils.

Number of origin groups

Because neither Table 3a nor Table 3b show any significant associations between the residualized number of origin groups and math performance, we reject the *educational*

Table 3a. The effects of the origin composition on math score of native students.

	Grade 2			Grade 5			Grade 8		
	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4
Constant	61.8**	61.8**	61.8**	79.3**	79.3**	79.3**	121.5**	121.5**	121.5**
	(0.6)	(0.6)	(0.6)	(0.7)	(0.7)	(0.7)	(0.5)	(0.5)	(0.5)
Class level									
Proportion migrants of class	–4.0**	–4.0**	–4.1*	–5.5**	–5.5**	–5.5**	–3.0**	–2.9**	–2.9**
	(1.1)	(1.1)	(1.1)	(1.2)	(1.2)	(1.2)	(0.9)	(0.9	(0.9)
Residualized origin diversity of class		–1.9			5.4			–9.2	
		(6.6)			(9.7)			(6.7)	
Proportion parents with tertiary education of class	0.1	0.1	0.1	1.3	1.3	1.3	1.1	1.2	1.1
	(1.2)	(1.2)	(1.2)	(1.4)	(1.4)	(1.4)	(1.1)	(1.1)	(1.1)
Residualized number of origin groups			0.2			0.1			–0.1
			(0.3)			(0.3)			(0.2)
Individual characteristics									
Parental education missing	–5.6**	–5.6**	–5.6**	–6.3**	–6.3**	–6.3**	–3.6**	–3.6**	–3.6**
	(1.6)	(1.6)	(1.6)	(1.5)	(1.5)	(1.5)	(1.0)	(1.0)	(1.0)
Low parental education	–8.1**	–8.1**	–8.0**	–12.9**	–12.9**	–12.9**	–6.7**	–6.8**	–6.7**
	(1.1)	(1.1)	(1.1)	(1.3)	(1.3)	(1.3)	(0.9)	(0.9)	(0.9)
Lower secondary parental education	–6.6**	–6.6**	–6.6**	–9.9**	–9.9**	–9.9**	–6.4**	–6.4**	–6.4**
	(0.4)	(0.4)	(0.4)	(0.5)	(0.5)	(0.5)	(0.3)	(0.3)	(0.3)
Upper secondary parental education	–3.6**	–3.6**	–3.6**	–4.8**	–4.8**	–4.8**	–3.4**	–3.4**	–3.4**
	(0.3)	(0.3)	(0.3)	(0.4)	(0.4)	(0.4)	(0.2)	(0.2)	(0.2)
Tertiary parental education	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Female	0.5	0.5	0.5	–5.3**	–5.3**	–5.3**	–2.8**	–2.8**	–2.8**
	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.2)	(0.2)	(0.2)
Variance									
Individual level	133.1**	133.1**	133.1**	174.6**	174.6**	174.6**	67.4**	67.4**	67.4**
	(2.1)	(2.1)	(2.1)	(2.8)	(2.8)	(2.8)	(1.2)	(1.2)	(1.2)
Class level	13.4**	13.4**	13.4**	17.8**	17.8**	18.0**	3.3**	3.1**	3.2**
	(1.9)	(1.9)	(1.9)	(3.3)	(3.3)	(3.3)	(1.2)	(1.2)	(1.2)

(*Continued*)

Table 3a. Continued.

	Grade 2			Grade 5			Grade 8		
	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4
School level	16.6**	16.6**	16.6**	9.6**	9.5**	9.4**	12.5**	12.5**	12.4**
	(2.4)	(2.4)	(2.4)	(3.2)	(3.2)	(3.2)	(1.7)	(1.7)	(1.7)
Log likelihood	68958.7	68958.6	68958.3	66979.0	66978.7	66978.8	52348.6	52346.7	52348.5
N students	8803			8290			7330		
N classes	991			677			580		
N schools	494			495			477		

Notes: Standard errors between brackets. Significance: $^{**}p < 0.01$; $^{*}p < 0.05$.
Source: COOL 2008, own computation.

Table 3b. The effects of the origin composition on math score of migrant students.

	Grade 2			Grade 5			Grade 8		
	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4
Constant	52.6**	52.6**	52.6**	69.4**	69.3**	69.3**	114.7**	114.7**	114.8**
	(0.9)	(0.9)	(1.2)	(1.7)	(1.7)	(1.7)	(1.1)	(1.1)	(1.1)
School level									
Proportion migrants	–1.6	–1.6	1.6	–3.8*	–3.7*	–3.7*	1.1	1.2	1.1
of class	(1.2)	(1.2)	(1.2)	(1.6)	(1.6)	(1.6)	(1.0)	(1.0)	(1.0)
Residualized origin diversity of class		–3.4			–5.0			–3.6	
		(3.8)			(5.0)			(4.0)	
Proportion parents with tertiary education of	4.0*	4.2*	4.1*	0.6	0.7	0.6	3.6	3.7	3.6
class	(1.8)	(1.8)	(1.8)	(2.6)	(2.6)	(2.6)	(1.9)	(1.9)	(1.9)
Residualized number of origin groups			–0.1			0.2			0.1
			(0.3)			(0.3)			(0.2)
Individual characteristics									
Parental education missing	–3.1	–3.1	–3.1	–4.2	–4.2	–4.2	–4.3**	–4.3**	–4.3**
	(2.0)	(2.0)	(2.0)	(2.2)	(2.2)	(2.2)	(1.4)	(1.4)	(1.4)
Low parental education	–4.2**	–4.2**	–4.2**	–5.0**	–5.0**	–5.0**	–3.6**	–3.6**	–3.6**
	(0.6)	(0.6)	(0.6)	(0.9)	(0.9)	(0.9)	(0.6)	(0.6)	(0.6)
Lower secondary parental education	–3.9**	–3.9**	–3.9**	–5.3**	–5.2**	–5.3**	–3.5**	–3.5**	–3.5**
	(0.6)	(0.6)	(0.6)	(0.9)	(0.9)	(0.9)	(0.6)	(0.6)	(0.6)
Upper secondary parental education	–1.9**	–1.9**	–1.9**	–2.4**	–2.4**	–2.4**	–1.8**	–1.8**	–1.8**
	(0.6)	(0.6)	(0.6)	(0.9)	(0.9)	(0.9)	(0.6)	(0.6)	(0.6)
Tertiary parental education	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Female	–0.0	–0.0	–0.0	–5.6**	–5.7**	–5.7**	–2.4**	–2.4**	–2.4**
	(0.4)	(0.4)	(0.4)	(0.5)	(0.5)	(0.5)	(0.3)	(0.3)	(0.3)
First-generation migrant	–0.9	–0.9	–0.9	–1.3	–1.4	–1.3	–1.6**	–1.6**	–1.6**
	(0.9)	(0.9)	(0.9)	(1.0)	(1.0)	(1.0)	(0.5)	(0.5)	(0.5)
Turkish origin	–0.8	–0.8	–0.8	3.1**	3.1**	3.1**	2.5**	2.5**	2.5**
	(0.7)	(0.7)	(0.7)	(1.0)	(1.0)	(1.0)	(0.6)	(0.6)	(0.6)

(*Continued*)

Table 3b. Continued.

	Grade 2			Grade 5			Grade 8		
	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4
Moroccan origin	–0.4	–0.4	–0.4	2.8**	2.8**	2.8**	1.9**	1.9**	1.9**
	(0.7)	(0.7)	(0.7)	(1.0)	(1.0)	(1.0)	(0.6)	(0.6)	(0.6)
Western origin	4.9**	4.9**	4.9**	5.7**	5.7**	5.7**	2.2*	2.1*	2.2*
	(1.0)	(1.0)	(1.0)	(1.4)	(1.4)	(1.4)	(0.9	(0.9	(0.9
Eastern-European origin	–0.2	–0.2	–0.2	5.3**	5.3**	5.3**	3.0**	3.0**	3.0**
	(1.1)	(1.1)	(1.1)	(1.5)	(1.5)	(1.5)	(0.9)	(0.9)	(0.9)
Chinese origin	5.2**	5.2**	5.2**	12.6**	12.6**	12.6**	8.3**	8.4**	8.3**
	(1.8)	(1.8)	(1.8)	(2.4)	(2.4)	(2.4)	(1.5)	(1.5)	(1.5)
Iraqi origin	–0.1	–0.1	–0.1	4.4*	4.4*	4.3*	5.6**	5.5**	5.5**
	(1.3)	(1.3)	(1.3)	(2.1)	(2.1)	(2.1)	(1.3)	(1.2)	(1.2)
Afghan origin	0.8	0.8	0.8	5.4*	5.4*	5.4*	5.6**	5.7**	5.6**
	(1.6)	(1.6)	(1.6)	(2.6)	(2.6)	(2.6)	(1.3)	(1.3)	(1.3)
Somali origin	1.6	1.6	1.6	4.5*	4.5*	4.5*	2.0	2.0	2.0
	(1.3)	(1.3)	(1.3)	(2.1)	(2.1)	(2.1)	(1.6)	(1.6)	(1.6)
Other country origin	0.9	0.9	0.9	3.7**	3.7**	3.7**	3.5**	3.5**	3.5**
	(0.7)	(0.7)	(0.7)	(1.0)	(1.0)	(1.0)	(0.6)	(0.6)	(0.6)
Former colony origin	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Variance									
Individual level	101.0**	101.0**	101.0**	190.6**	190.6**	190.7**	71.8**	71.8**	71.8**
	(2.7)	(2.7)	(2.7)	(5.2)	(5.2)	(5.2)	(2.1)	(2.1)	(2.1)
Class level	17.4**	17.4**	17.4**	10.6*	10.7*	10.8*	1.7	1.9	1.6
	(3.0)	(3.0)	(3.0)	(4.8)	(4.8)	(4.8)	(1.7)	(1.8)	(1.7)
School level	14.9**	14.8**	14.9**	23.7**	23.7**	23.4**	11.7**	11.3**	11.8**
	(3.2)	(3.2)	(3.2)	(5.9)	(5.9)	(5.9)	(2.4)	(2.3)	(2.3)
Log likelihood	26090.2	26089.3	26090.1	25020.7	25019.7	25020.3	19611.5	19610.7	19611.4
N students	3426			3056			2719		
N classes	719			505			438		
N schools	369			383			362		

Notes: Standard errors between brackets. Significance: $^{**}p < 0.01$; $^{*}p < 0.05$.

Source: COOL 2008, own computation.

Table 4a. The effects of the origin composition on reading comprehension of students with a native background[a].

	Grade 2			Grade 5			Grade 8		
	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4
Constant	75.8**	74.9**	75.9**	30.9**	30.9**	30.9**	60.6**	60.6**	60.6**
	(0.5)	(0.3)	(0.5)	(0.6)	(0.6)	(0.6)	(0.8)	(0.8)	(0.8)
School level									
Proportion migrants	–5.1**	–5.1**	–5.2**	–4.0**	–4.0**	–4.0**	–3.4**	–3.2**	–3.2**
of class	(0.8)	(0.8)	(0.8)	(1.1)	(1.1)	(1.1)	(1.3)	(1.3)	(1.3)
Residualized origin diversity of class		–1.4			–1.6			–15.9	
		(5.0)			(9.0)			(9.9)	
Proportion parents with tertiary education	0.4	0.4	0.4	2.1	2.1	2.1	4.5**	4.6**	4.8**
of class	(0.9)	(0.9)	(0.9)	(1.2)	(1.2)	(1.2)	(1.6)	(1.6)	(1.6)
Residualized number			0.1			–0.1			–0.6*
of origin groups			(0.2)			(0.3)			(0.3)
Individual characteristics									
Parental education missing	–2.5*	–2.5*	–2.5*	–6.1**	–6.1**	–6.1**	–5.8**	–5.8**	–5.9**
	(1.2)	(1.2)	(1.2)	(1.4)	(1.4)	(1.4)	(1.6)	(1.6)	(1.6)
Low parental education	–6.4**	–6.4**	–6.4**	–12.9**	–12.9**	–12.9**	–15.2**	–15.3**	–15.2**
	(0.8)	(0.8)	(0.8)	(1.3)	(1.3)	(1.3)	(1.5)	(1.5)	(1.5)
Lower secondary parental education	–4.4**	–4.4**	–4.5**	–9.3**	–9.3**	–9.3**	13.5**	13.5**	13.5**
	(0.3)	(0.3)	(0.3)	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)
Upper secondary parental education	–2.1**	–2.1**	–2.1**	–5.7**	–5.7**	–5.7**	–7.4**	–7.4**	–7.4**
	(0.2)	(0.2)	(0.2)	(0.4)	(0.4)	(0.4)	(0.4)	(0.4)	(0.4)
Tertiary parental education	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Female	1.3**	1.3**	1.3**	2.9**	2.9**	2.9**	3.0**	3.0**	3.0**
	(0.2)	(0.2)	(0.2)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)
Variance									
Individual level	70.2**	70.2**	70.2**	162.8**	162.8**	162.8**	200.4**	200.3**	200.4**
	(1.1)	(1.1)	(1.1)	(2.7)	(2.7)	(2.7)	(3.4)	(3.4)	(3.4)
Class level	8.5**	8.5**	8.5**	7.5**	7.5**	7.5**	18.8**	18.6**	18.4**
	(1.1)	(1.1)	(1.1)	(2.1)	(2.1)	(2.1)	(4.3)	(4.3)	(4.3)

(*Continued*)

Table 4a. Continued.

	Grade 2			Grade 5			Grade 8		
	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4
School level	11.5**	11.5**	11.6**	5.6**	5.6**	5.6**	6.0	6.0	6.1
	(1.5)	(1.5)	(1.5)	(2.0)	(2.0)	(2.0)	(4.0)	(3.9)	(3.9)
Log likelihood	63621.8	63621.7	63621.3	64229.2	64229.2	64229.0	62060.2	62060.2	62058.5
N students	8826			8047			7563		
N classes	990			670			590		
N schools	496			490			481		

Notes: [a]For Grade 2, the reading comprehension test is not available. We used the Cito language score of Grade 2 instead. Because the language score contains both oral communication and early literacy, the results of Grade 2 are not directly comparable to those of Grades 5 and 8. Standard errors between brackets. Significance: $^{**}p < 0.01$; $^{*}p < 0.05$.

Source: COOL 2008, own computation.

Table 4b. The effects of the origin composition on reading comprehension of students with a migrant background.

	Grade 2			Grade 5			Grade 8		
	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4
Constant	69.0**	68.9**	69.0**	22.7**	22.6**	22.7**	52.8**	52.7**	52.8**
	(0.9)	(0.9)	(0.9)	(1.3)	(1.3)	(1.3)	(1.6)	(1.6)	(1.6)
School level									
Proportion migrants	–1.1	–1.0	–1.1	–0.9	–0.9	–1.0	–2.1	–1.9	–2.0
of class	(0.9)	(0.9)	(0.9)	(1.1)	(1.1)	(1.1)	(1.5)	(1.5)	(1.5)
Residualized origin diversity of class		–3.7			–3.3			–12.7*	
		(2.8)			(3.1)			(5.3)	
Proportion parents with tertiary education	2.7*	2.8*	2.7*	5.1**	5.1**	5.1**	5.4*	5.5*	5.6*
of class	(1.4)	(1.4)	(1.4)	(1.8)	(1.8)	(1.8)	(2.6)	(2.6)	(2.6)
Residualized number			–0.3			–0.3			–0.3
of origin groups			(0.2)			(0.2)			(0.3)
Individual characteristics									
Parental education missing	–3.0	–3.0	–2.9	–5.0**	–5.0**	–5.0**	–7.3**	–7.3**	–7.3**
	(1.6)	(1.6)	(1.6)	(1.8)	(1.8)	(1.8)	(2.1)	(2.1)	(2.1)
Low parental education	–3.5**	–3.5**	–3.5**	–4.8**	–4.8**	–4.8**	–6.9**	–6.8**	–6.9**
	(0.5)	(0.5)	(0.5)	(0.8)	(0.8)	(0.8)	(0.9)	(0.9)	(0.9)
Lower secondary parental education	–2.2**	–2.2**	–2.2**	–4.8**	–4.8**	–4.8**	–7.1**	–7.1**	–7.1**
	(0.5)	(0.5)	(0.5)	(0.8)	(0.8)	(0.8)	(0.9)	(0.9)	(0.9)
Upper secondary parental education	–1.0*	–1.0*	–1.0*	–2.4**	–2.4**	–2.4**	–3.2**	–3.2**	–3.2**
	(0.4)	(0.5)	(0.4)	(0.7)	(0.7)	(0.7)	(0.9)	(0.9)	(0.9)
Tertiary parental education	Ref.	Ref.	Ref.	Ref	Ref	Ref	Ref	Ref	Ref
Female	0.9**	0.9**	0.9**	2.3**	2.3**	2.3**	2.6**	2.6**	2.6**
	(0.3)	(0.3)	(0.3)	(0.4)	(0.4)	(0.4)	(0.5)	(0.5)	(0.5)
First-generation migrant	–1.6*	–1.6*	–1.6*	–1.3	–1.3	–1.3	–1.4	–1.4	–1.4
	(0.7)	(0.7)	(0.7)	(0.8)	(0.8)	(0.8)	(0.8)	(0.8)	(0.8)
Turkish origin	–2.6**	–2.6**	–2.6**	–2.1*	–2.0*	–2.1*	–2.2*	–2.2*	–2.2*
	(0.5)	(0.5)	(0.5)	(0.8)	(0.8)	(0.8)	(0.9	(0.9	(0.9
Moroccan origin	–0.7	–0.8	–0.8	–0.5	–0.5	–0.6	0.5	0.4	0.4
	(0.5)	(0.5)	(0.5)	(0.8)	(0.8)	(0.8)	(1.0)	(0.9)	(1.0)

(*Continued*)

Table 4b. Continued.

	Grade 2			Grade 5			Grade 8		
	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4
Western origin	3.3**	3.3**	3.3**	2.3*	2.3*	2.2	2.5	2.5	2.5
	(0.8)	(0.8)	(0.8)	(1.2)	(1.2)	(1.2)	(1.4)	(1.4)	(1.4)
Eastern-European origin	–1.8*	–1.7*	–1.7*	–0.1	–0.1	–0.1	3.2*	3.2*	3.2*
	(0.8)	(0.8)	(0.8)	(1.2)	(1.2)	(1.2)	(1.4)	(1.4)	(1.4)
Chinese origin	–1.5	–1.5	–1.5	5.0**	5.1**	5.1**	4.4*	4.5*	4.5*
	(1.4)	(1.4)	(1.4)	(1.9)	(1.9)	(1.9)	(2.2)	(2.3)	(2.2)
Iraqi origin	–1.5	–1.4	–1.5	–1.7	–1.7	–1.7	0.3	0.4	0.4
	(1.0)	(1.0)	(1.0)	(1.7)	(1.7)	(1.7)	(1.8)	(1.8)	(1.8)
Afghan origin	–0.7	–0.7	–0.7	1.7	1.7	1.7	5.8**	5.9**	5.8**
	(1.2)	(1.2)	(1.2)	(2.1)	(2.1)	(2.1)	(2.0)	(2.0)	(2.0)
Somali origin	0.0	0.1	0.1	2.6	2.7	2.6	–0.5	–0.5	–0.5
	(1.0)	(1.0)	(1.0)	(1.7)	(1.7)	(1.7)	(2.3)	(2.3)	(2.3)
Other country origin	–0.5	–0.5	–0.5	2.5**	2.5**	2.4**	3.0**	3.0**	3.0**
	(0.5)	(0.5)	(0.5)	(0.8)	(0.8)	(0.8)	(1.0)	(1.0)	(1.0)
Former colony origin	Ref.	Ref.	Ref.	Ref	Ref	Ref	Ref	Ref	Ref
Variance									
Individual level	60.0**	60.0**	60.0**	123.7**	123.8**	123.7**	176.5**	176.6**	176.5**
	(1.6)	(1.6)	(1.6)	(3.6)	(3.6)	(3.6)	(5.0)	(5.0)	(5.0)
Class level	7.2**	7.3**	7.4**	3.1	3.1	3.0	12.8**	14.7**	13.2**
	(1.5)	(1.5)	(1.5)	(2.6)	(2.6)	(2.6)	(5.5)	(5.7)	(5.6)
School level	12.4**	12.1**	12.0**	5.8*	5.7*	5.9*	9.8	6.5	8.9
	(2.1)	(2.0)	(2.0)	(2.6)	(2.6)	(2.6)	(5.3)	(5.3)	(5.4)
Log likelihood	24392.6	24390.9	24390.3	21209.1	21208.0	21207.3	22669.0	22663.5	22668.1
N students	3441			2751			2801		
N classes	720			492			452		
N schools	371			374			368		

Notes: Standard errors between brackets. Significance: $^{**}p < 0.01$; $^{*}p < 0.05$.
Source: COOL 2008, own computation.

instruction hypothesis (see Driessen, 2002, for a similar result). A higher number of origin groups in a class than is typical given the proportion of migrants is not related to lower math scores. Nevertheless, if we look at reading comprehension, Table 4a shows for the native pupils in Grade 8 a significant effect of –0.6 of the number of origin groups on reading comprehension. Therefore, we only found evidence in favour of the *educational instruction hypothesis* for the reading comprehension for the native pupils in Grade 8.[8]

Inspection of the log likelihood ratio in Tables 3a, 3b, 4a, and 4b learns that including the number of origin groups residuals (Model 4) led to an improved model fit compared to a model without the number of origin groups (Model 2).

Origin diversity residuals

Model 3 in Table 3a shows parameter estimates of residualized origin diversity of –1.9 in Grade 2, 5.4 in Grade 5, and finally –9.2 in Grade 8 (all for natives). However, none of these negative coefficients is statistically significant at $p < 0.05$. For migrant pupils (Table 3b), origin diversity is always negatively related to mathematics achievement, although in none of the models in a statistically significant way. Given these non-significant results, the origin diversity hypothesis is rejected with regard to mathematics.

Tables 4a and 4b turn the attention to reading comprehension as the dependent variable. For native pupils, we found no significant associations between residualized diversity and reading comprehension, again refuting the origin diversity hypothesis. Migrant children, however, are negatively affected by origin diversity, in particular in Grade 8 (regression coefficient of –12.7). Consequently, in Grade 8 we see that children of migrant descent have lower performance in reading in school classes with an origin diversity that is higher than we expect on the basis of the number of migrants in a class. We therefore find evidence in favour of the *origin diversity hypothesis* in Grade 8 for reading comprehension for the migrants.

The log likelihood ratio shows in most cases a better fit when we include origin diversity residuals. This, however, does not hold for reading comprehension of natives in Grades 5 and 8.

Should we measure origin composition at the class or at the school grade level?

On theoretical grounds, we employed a three-level model (individual, class, school grade) with compositional measures taken at the class level. In most cases, the variance at the class level (within school grades) was significant. Here, we demonstrate whether results are different when the class level was omitted from the research design and compositional measures were instead taken at the school grade level (as has been done in previous research). Appendix 4 shows the tables of the two-level multilevel analyses. The following differences were encountered.

First, in contrast to our results in Table 4a, for the natives in Grade 8 no significant effect of the residualized number of origin groups (at the school grade level) on reading comprehension was found. Also in our original model, the coefficient was small (but significant).

Second, unlike the analyses described above, we found significant negative effects of residualized origin diversity for native students in Grade 2, and significant negative effects of the number of origin groups for natives in Grades 2 and 5.

Third, for migrant students the two-level model with measures at the school grade level also shows a significant effect of residualized origin diversity in Grade 2. For the migrant students in Grades 2 and 5, significant negative effects were found of the number of origin groups on reading comprehension scores.

In sum, examining compositional measures at the school level (separately by grade) shows slightly stronger effects of diversity indicators on school performance. The discrepancy in the findings may be related to schools having larger diversities within grades than within classes within grades.

Conclusions and discussion

We empirically explored the association between various indicators of the origin composition of school classes on pupils' test scores in mathematics and reading comprehension at different grades in primary school in The Netherlands. Our particular interest was in the associations between academic performance and two distinct characteristics on origin compositions of school classes: the *proportion* of migrants (first- and second-generation) and the *diversity* among the different origin groups.

In studying diversity, a further distinction was made between the number of origin groups in a class and a diversity index which includes information about the number of groups and the sizes of the groups jointly. Both variables have been residualized on the proportion of migrants in a class, implying that these diversity indicators measure the relative diversity given a particular proportion of first- and second-generation ethnic minority children. The diversity index has been associated with a combination of instructional problems and peer group effects, whereas the number of origin groups more clearly relates to instructional problems of diverse classes.

Our results demonstrated that the proportion of migrants in a class is negatively related to academic performance of native pupils. Pupils of migrant origins are less strongly affected by larger proportions of migrants in a class. The diversity of pupils in terms of origin has weaker effects overall, although the reading comprehension of children with a migration background is negatively related to origin diversity in Grade 8, the year in which decisions are made for the school type that can be attended in secondary education. This conforms to the study of Van Ewijk and Sleegers (2010a), who demonstrated that peer-group effects increase as pupils get older.

Also children of Dutch descent had slightly lower scores on reading comprehension if they were in a class with a larger number of different origin groups. So, instructional problems resulting from a larger number of origin groups in a class were more negatively affecting children of Dutch descent. The combined results suggest that for reading comprehension in Grade 8 native pupils are significantly influenced by *instructional mechanisms* and migrant pupils by a combination of *instructional and peer group mechanisms*.

The difference between math and reading comprehension for the effect of the number of origin groups for native pupils in Grade 8 could possibly be caused by the different instructional needs of the pupils with respect to math and reading comprehension. For instance, for reading comprehension native pupils possibly need more instruction that connects to their own needs as natives. The results suggest that teachers have more problems with also planning the instruction for the needs of the native pupils. Nevertheless, for math the teachers could possibly instruct more origin groups at the same time.

Our findings of origin diversity residuals in Grade 8 are partially in line with the earlier research of Dronkers and Van der Velden (in press). Using data from diverse OECD countries, Dronkers and Van der Velden found significant negative effects of origin diversity on reading scores for migrants in secondary schools. Nevertheless, in contrast to this earlier study we found in our research model no significant effects for mathematics. This difference could possibly be caused by the difference between primary and secondary education and the use of more countries in the analysis of Dronkers and Van der Velden.

Furthermore, the research design is different because we used (methodologically preferred) residualized scores on diversity. In our view, such a solution to the multicollinearity problem should be addressed in future research.

A previous study, using older Dutch primary school data (*Primair onderwijs en speciaal onderwijs cohortonderzoeken,* PRIMA), showed positive effects in higher grades of origin diversity on math scores (Maestri, 2011a). This study used another research design with grade means. Nevertheless, the recent COOL data could give a clarification of the differences in the earlier findings of the effect of origin diversity. A possible clarification for this finding can be that the dataset of the earlier study included a more restricted collection of origin groups in the classrooms. Although both in the PRIMA and COOL data the origin countries are classified in 15 categories, in the PRIMA data some origin countries are old European migration countries like Spain, Italy, Greece, and Portugal, and in the COOL data these origin categories are replaced by new migration countries like Iraq, Afghanistan, Somalia, and a new European migration country: Poland. Therefore, a stronger origin diversity in the older PRIMA data relates to different, mostly European and Mediterranean, migrants in the classroom. In this recent study, however, a stronger origin diversity relates to pupils from highly diverse new migrant countries. Therefore, we could expect that due to the composition of the cultural distances with a higher origin diversity in the COOL data, this diversity could lead to more problems in ethnic identification and interethnic conflicts than in the PRIMA origin diversity index. Consequently, future research could take into account the difference in the influence of the cultural distances in the origin diversity index between diverse ethnic groups.

It must be noted that, with the cross-sectional data that we (and others) have used, claims about causal effects of class composition on pupils' performance should be made with caution. In particular, given the free school choice policy in The Netherlands, it is possible that native families are more concerned about the ethnic composition of schools than migrant families. If better performing native pupils are more likely to avoid schools with large concentrations of migrant children, for example, because their better educated parents are better informed or more concerned, it is possible that our observed negative relationship between the proportion of migrants' and natives' reading comprehension is flawed by this school selection process.

Although this study offers an expansion on the earlier models that measure the origin composition, future research can enrich these findings by using forthcoming cohort data from COOL, secondary school data, data from other countries, separate analyses for different origin groups, and an analysis for other non-cognitive school outcomes like active citizenship.

Acknowledgement

The authors would like to thank all reviewers for their helpful comments. Gert-Jan M. Veerman received support from the Doctoral Grant for Teachers funded by the Netherlands Organization for Scientific Research [grant number 023.001.120].

Notes

1. Van Houtte and Stevens (2009) also focused on the distinction between ethnic share and ethnic diversity; however, they measured the effect on the sense of belonging in school.
2. Although earlier research consequently uses the concept ethnicity, we prefer the concept origin country (Hutchinson & Smith, 1996). Our data contain, for example, for pupils with parents of Turkish origin only information about their origin country. However, these parents could have, for example, a Turkish, a Kurdish, or an Armenian ethnicity.

3. NWO grant numbers 411.20.411 and 411.20.412.
4. Dutch pupils enrol in Grade 1 at the day they turn 4 years old. Grades 1–2 are comparable to Kindergarten in many other systems, and the final Grade 8 of primary school thus equates to Grade 6 in many other systems.
5. In Grade 5, schools could use two different versions of the Cito math test. The two versions were made comparable using conversion tables made by Cito.
6. $VIF = -\frac{1}{1 - R_i^2}$
7. VIF statistics are possible using OLS; nevertheless, VIF's are not available in multilevel analysis due to the lack of R^2. Because we have different levels in the other models, VIF's are not available for the other models. Due to the low correlation of the other variables, we know that multicollinearity is only somewhat stronger in the multilevel models with more variables.
8. We also tested the effect of origin diversity residuals and of the number of origin groups residuals for the students on reading vocabulary in Grades 5 and 8. The results show also a significant negative effect for the number of origin groups residuals for natives in Grade 8. However, the results show no significant effects for origin diversity residuals and the number of origin groups residuals in the other grades for both natives and migrants. Therefore, we also reject the teaching hypothesis for the natives in Grade 8 for reading vocabulary. Results available on request.

Notes on contributors

Gert-Jan M. Veerman was a teacher at primary schools in The Netherlands for more than 10 years. He is currently lecturer Education at the *Ede Christian University of Applied Sciences* and PhD student at the *University of Amsterdam*, The Netherlands. In 2011, he received the Doctoral Grant for Teachers from the Netherlands Organization for Scientific Research (NWO). He holds a MSc in sociology and a MSc in education from the *University of Amsterdam*. His research interests include ethnicity and school performances.

Herman G. van de Werfhorst is professor of sociology at the University of Amsterdam, and is director of the Amsterdam Centre for Inequality Studies (AMCIS). His prime research interests are in the comparison of educational systems with regard to social and ethnic selection, labour market outcomes, and civic engagement. His work has appeared in many academic journals, including *Sociology of Education*, the *European Sociological Review*, the *British Educational Research Journal*, the *Journal of Education and Work*, the *British Journal of Sociology*, the *British Journal of Sociology of Education*, and *Demography*.

Jaap Dronkers studied sociology at the *Vrije Universiteit Amsterdam.* He worked at the *Catholic University Brabant* as associate professor and from 1990 at the University of Amsterdam, first as professor in the educational sciences, later as professor of empirical sociology. From 2001 until 2009, he was Professor of Social Stratification and Inequality at the *European University Institute* (EUI) in Florence. From December 2009, he holds the chair international comparative research on educational performance and social inequality at the *Maastricht University.* He has published on the causes and consequences of unequal educational and occupational attainment, changes in educational opportunities, effect differences between public and religious schools, the educational and occupational achievement of migrants from different origins and in various countries of destination, the linkages between school and the labour market, the effects of parental divorce on children, cross-national differences in causes of divorce, education of Dutch elites, and European nobility. He also participates in public debates on topics related to his research.

References

Agirdag, O., Van Houtte, M., & Van Avermaet, P. (2012). Why does the ethnic and socio-economic composition of schools influence math achievement? The role of sense of futility and futility culture. *European Sociological Review*, *28*, 366–378. doi: 10.1093/esr/jcq070

Creemers, B. P. M. (2007). Educational effectiveness and improvement: The development of the field in mainland Europe. In T. Townsend (Ed.), *International handbook of school effectiveness and improvement* (pp. 223–242). New York, NY: Springer.

Demanet, J., Agirdag, O., & Van Houtte, M. (2011). Etnische schoolsamenstelling en sociaal kapitaal: Een onderzoek naar de effecten van etnische concentratie en heterogeniteit op verbondenheid met vrienden, leerkrachten en de school bij autochtone en allochtone leerlingen [Ethnic school composition and social capital: A study on the effects of ethnic concentration and heterogeneity of connectedness with friends, teachers and the school for native and immigrant students]. In J. Dronkers (Ed.), *Goede bedoelingen in het onderwijs: Kansen en missers* [Good intentions in education: Opportunities and misses] (pp. 105–132). Amsterdam, The Netherlands: Amsterdam University Press.

Driessen, G. (2002). School composition and achievement in primary education: A large-scale multilevel approach. *Studies in educational evaluation*, *28*, 347–368. Retrieved from http://www.its-nijmegen.nl/medewerkers/drig/pdf/see02-driessen-schoolcomposition.pdf

Driessen, G., Mulder, L., Ledoux, G., Roeleveld, J., & Van der Veen, I. (2009). *Cohortonderzoek COOL5-18. Technisch rapport basisonderwijs, eerste meting 2007/08* [Cohort study COOL5-18. Technical report primary education, first measurement 2007/08]. Nijmegen/Amsterdam, The Netherlands: ITS/SCO-Kohnstamm Instituut.

Driessen, G., Mulder, L., & Roeleveld, J. (2012). *Cohortonderzoek COOL5-18. Technisch rapport basisonderwijs, tweede meting 2010/11* [Cohort study COOL5-18. Technical report primary education, second measurement 2010/11]. Nijmegen/Amsterdam, The Netherlands: ITS/SCO-Kohnstamm Instituut.

Dronkers, J. (2004). Do public and religious schools really differ? Assessing the European evidence. In P. J. Wolf & S. Macedo (Eds.), *Educating citizens. International perspectives on civic values and school choice* (pp. 287–314). Washington, DC: Brookings Institution Press.

Dronkers, J., & Van der Velden, R. (in press). Positive but also negative effects of ethnic diversity in schools on educational performance? An empirical test using cross-national PISA data. In M. Windzio (Ed.), *Integration and inequality in educational institutions*. Dordrecht, The Netherlands: Springer.

Gijsberts, M., Van der Meer, T., & Dagevos, J. (2012). "Hunkering down" in multi-ethnic neighbourhoods? The effects of ethnic diversity in cities and neighbourhoods on dimensions of social cohesion. *European Sociological Review*, *28*, 527–537. doi:10.1093/esr/jcr022

Herweijer, L. (2011). *Gemengd leren: Etnische diversiteit en leerprestaties in het basis- en voortgezet onderwijs* [Mixed learning: Ethnic diversity and learning in primary and secondary education]. Den Haag, The Netherlands: SCP.

Hoxby, C. (2000). *Peer effects in the classroom: Learning from gender and race variation* (NBER Working Paper No. 7867). Retrieved from http://www.nber.org/papers/w7867

Hutchinson, J., & Smith, A. D. (Eds.). (1996). *Ethnicity.* Oxford, UK: Oxford University Press.

Karsten, S. (2010). School segregation. In Organisation of Economic Co-operation and Development (Ed.), *Equal opportunities? The labour market integration of the children of immigrants* (pp. 193–209). Paris. France: Organisation of Economic Co-operation and Development.

Kleinbaum, D. G., Kupper, L. L., Nizam, A., & Muller, K. E. (2008). *Applied regression analysis and other multivariable methods* (4th ed.). Belmont, CA: Thomson Brooks/Cole.

Ladd, H. F., & Fiske, E. B. (2009). *The Dutch experience with weighted student funding: Some lessons for the U.S.* (Sanford Working Papers Series SAN09-03). Retrieved from http://sanford.duke.edu/research/papers/SAN09-03.pdf

Ladd, H. F., Fiske, E. B., & Ruijs, N. (2010). *Parental choice in the Netherlands: Growing concerns about segregation* (Sanford Working Papers Series SAN10-02). Retrieved from http://sanford.duke.edu/research/papers/SAN10-02.pdf

Lancee, B., & Dronkers, J. (2011). Ethnic, religious and economic diversity in Dutch neighbourhoods: Explaining quality of contact with neighbours, trust in the neighbourhood and inter-ethnic Trust. *Journal of Ethnic and Migration Studies*, *37*, 597–618. doi:10.1080/1369183X.2011.545277

Lazear, E. P. (1998). *Diversity and immigration* (NBER Working Paper No. 6535). Retrieved from http://www.nber.org/papers/w6535

Lazear, E. P. (1999). Culture and language. *Journal of Political Economy*, *107*(56), s95–s126. doi:10.1086/250105

Levels, M., & Dronkers, J. (2008). Educational performance of native and immigrant children from various countries of origin. *Ethnic and Racial Studies*, *31*, 1404–1425. doi: 10.1080/01419870701682238

Lucassen, J., & Penninx, R. (1997). *Newcomers: Immigrants and their descendants in the Netherlands 1550–1995*. Amsterdam, The Netherlands: Het Spinhuis.

Maestri, V. (2011a). *A deeper insight into the ethnic make-up of school cohorts: Diversity and school achievement* (AIAS Working Paper 11–108). Amsterdam, The Netherlands: University of Amsterdam. Retrieved from http://www.uva-aias.net/uploaded_files/publications/WP108-Maestri.pdf

Maestri, V. (2011b). Een nadere beschouwing van de samenstelling van etnische minderheden op Nederlandse basisscholen. Diversiteit en leerprestaties [A further examination of the composition of ethnic minorities in Dutch primary schools. Diversity and learning performance]. In J. Dronkers (Ed.), *Goede bedoelingen in het onderwijs: Kansen en missers* [Good intentions in education: opportunities and misses] (pp. 155–178). Amsterdam, The Netherlands: Amsterdam University Press.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*, 673–690. doi:10.1007/s11135-006-9018-6

O'Reilly, C. A., III, Williams, K. Y., & Barsade, S. (1997). *Demography and group performance: Does diversity help?* (Stanford University Research Paper No. 1426). Retrieved from https://gsbapps.stanford.edu/researchpapers/library/RP1426.pdf

Organisation of Economic Co-operation and Development. (2011). *International migration outlook 2011*. Paris, France: OECD Publishing. doi:10.1787/migr_outlook-2011-en

Peetsma, T., Van der Veen, I., Koopman, P., & Van Schooten, E. (2006). Class composition influences on pupils' cognitive development. *School Effectiveness and School Improvement*, *17*, 275–302. doi:10.1080/13803610500480114

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectations and student intellectual development.* New York, NY: Holt, Rinehart & Winston.

Svalberg, A. M.-L. (2007). Language awareness and language learning. *Language Teaching*, *40*, 287–308. doi:10.1017/S0261444807004491

Tolsma, J., Van der Meer, T., & Gesthuizen, M. (2009). The impact of neighbourhood and municipality characteristics on social cohesion in The Netherlands. *Acta Politica*, *44*, 286–313. doi:10.1057/ap.2009.6

Van Ewijk, R. J. G., & Sleegers, P. (2010a). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, *5*, 134–150. doi:10.1016/j.edurev.2010.02.001

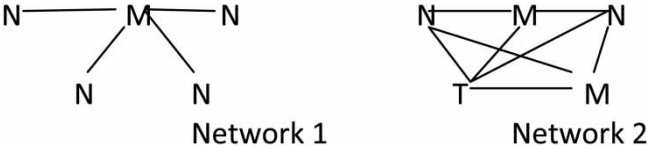
Van Ewijk, R. J. G., & Sleegers, P. (2010b). Peer ethnicity and achievement: A meta-analysis into the compositional effect. *School Effectiveness and School Improvement*, *21*, 237–265. doi:10.1080/09243451003612671

Van Houtte, M., & Stevens, P. A. J. (2009). School ethnic composition and students' integration outside and inside schools in Belgium. *Sociology of Education*, *82*, 217–239. doi:10.1177/003804070908200302

Appendix 1. Origin diversity, the number of origin groups and interethnic contacts

Although the number of origin groups and the origin diversity index are related, Figure 1.1 shows that a higher number of origin groups at class level does not necessarily lead to a higher diversity index. We computed for the native students a Pearson correlation of 0.90 for origin diversity and the number of origin groups; therefore, origin diversity explains 81% of the number of origin groups. Nevertheless, there are classes with both higher and lower numbers of origin groups than we would expect due to the origin diversity.

To show the impact of the origin composition on the indexes, we calculate the four indexes of origin composition with 10 classes with different origin compositions. Networks 1 and 2 below show what we mean by the number of possible interethnic contacts, using a group of 5 students, containing 1 Moroccan student and 4 native students. Moreover, Network 2 shows a group of 5 students, containing 2 native students, 2 Moroccan students, and 1 Turkish student. Network 1 leads to four possible interethnic contacts and Network 2 leads to eight possible interethnic contacts.



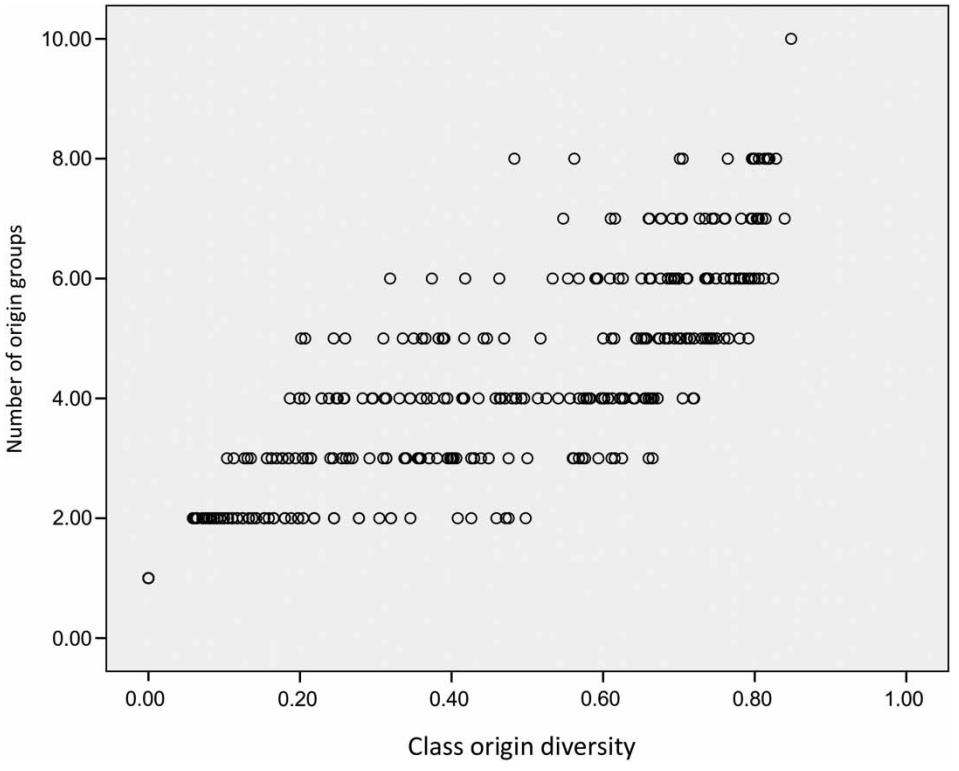


Figure 1.1. Class origin diversity and number of origin groups for native students in Grade 8. Source: COOL 2008, own computation.

As Table 1.1 shows, the number of students within the different origin groups influences the height of the origin diversity index. Therefore, a higher number of origin groups together with a maximum number of students within every ethnic group lead to the highest origin diversity score. Furthermore, Table 1.1 shows that a higher origin diversity index leads to a higher number of possible interethnic contacts within the group. Consequently, when we use the four indexes of the origin composition, origin diversity refers to the number of possible interethnic contacts within the class. Although the number of possible interethnic contacts is dependent on the group size, a higher origin diversity index refers to a higher number of possible interethnic group contacts within the possibilities of the group size. In contrast to the possible number of interethnic contacts within the group, the origin diversity index is not dependent on the group size; therefore, we could use the origin diversity index to compare the influence of the possible interethnic contacts for classes with different group sizes.

Table 1.1 Different ethnic composition scores with a school class of 20 students.

% migrants	Number of ethnic groups	Origin diversity	Possible interethnic contacts	Native	Moroccan	Turkish	Former Colonies
50	2	0.50	100	10	10	0	0
25	2	0.37	75	15	5	0	0
5	2	0.09	19	19	1	0	0
95	2	0.09	19	1	19	0	0
65	3	0.66	133	7	7	6	0
25	3	0.40	81	15	3	2	0
10	3	0.18	37	18	1	1	0
75	4	0.75	150	5	5	5	5
30	4	0.48	96	14	2	2	2
15	4	0.27	54	17	1	1	1

Appendix 2. Missing data

Table 2.1. Means for migrant students and native students with and without math scores.

	Grade 2				Grade 5				Grade 8			
	Migrants		Natives		Migrants		Natives		Migrants		Natives	
	Test scores	Test scores missing	Test scores	Test scores missing	Test scores	Test scores missing	Test scores	Test scores missing	Test scores	Test scores missing	Test scores	Test scores missing
% migrant students of class	63.22	63.25	14.37	14.31	63.67	65.24	13.35	13.97	62.75	67.06	13.90	13.97
Origin diversity of class	0.60	0.57	0.20	0.20	0.59	0.55	0.19	0.25	0.59	0.59	0.20	0.21
Number of origins in class	4.46	4.28	2.25	2.22	4.85	4.51	2.50	2.69	4.89	4.87	2.65	2.65
Parental education missing	0.01	0.02	0.02	0.02	0.02	0.02	0.03	0.05	0.02	0.03	0.02	0.07
Lower parental education	0.28	0.30	0.02	0.02	0.30	0.26	0.02	0.02	0.30	0.38	0.02	0.03
Lower secondary parental education	0.23	0.21	0.16	0.17	0.24	0.28	0.17	0.18	0.25	0.24	0.20	0.20
Upper secondary parental education	0.30	0.33	0.47	0.42	0.30	0.27	0.46	0.44	0.27	0.23	0.46	0.36
Tertiary parental education	0.18	0.14	0.33	0.37	0.15	0.16	0.32	0.33	0.15	0.12	0.30	0.34
Female	0.47	0.45	0.48	0.46	0.50	0.54	0.50	0.47	0.48	0.51	0.49	0.56
First-generation migrant	0.06	0.06			0.09	0.14			0.13	0.11		
Only student origin	0.33	0.34	0.01	0.01	0.25	0.29	0.004	0.003	0.26	0.31	0.003	0.004
Turkish origin	0.24	0.20			0.27	0.21			0.28	0.21		
Moroccan	0.24	0.29			0.25	0.24			0.22	0.27		
Western	0.07	0.07			0.06	0.07			0.06	0.04		
Eastern Europe	0.05	0.06			0.05	0.04			0.05	0.05		
Chinese	0.01	0.01			0.01	0.02			0.02	0.03		
Iraqi	0.03	0.02			0.02	0.02			0.02	0.02		
Afghan	0.02	0.02			0.01	0.01			0.02	0.02		
Somali	0.03	0.02			0.02	0.01			0.01	0.02		
Other country	0.20	0.19			0.18	0.19			0.16	0.18		

(*Continued*)

Table 2.1. Continued.

	Grade 2				Grade 5				Grade 8			
	Migrants		Natives		Migrants		Natives		Migrants		Natives	
	Test scores	Test scores missing	Test scores	Test scores missing	Test scores	Test scores missing	Test scores	Test scores missing	Test scores	Test scores missing	Test scores	Test scores missing
Former colony	0.12	0.12			0.13	0.19			0.15	0.17		
Math score	51.83		58.62		63.60		72.01		114.38		117.14	
N students	3462	311	9006	653	3082	281	8439	529	2742	377	7399	836
% missing test score		8%		7%		8%		6%		12%		10%

Source: COOL 2008, own computation.

Appendix 3. Residualized origin diversity

Figure 3.1 shows a linear and quadratic regression line for origin diversity as a function of origin share for the native students in Grade 8. The figure shows that we should explain origin diversity as a function of the proportion of migrants with a quadratic regression, because the quadratic line better predicts the origin diversity. The cases that lie independently from the quadratic line are the residuals. The residuals could be under the quadratic line and therefore negative and also above the line and therefore positive. A positive residual shows that the possible relative number of interethnic contacts is higher than the quadratic regression predicts with the given proportion of migrants. Furthermore, next to the figure we show some examples of the measurement of the distance of the origin diversity of the class from the regression line.

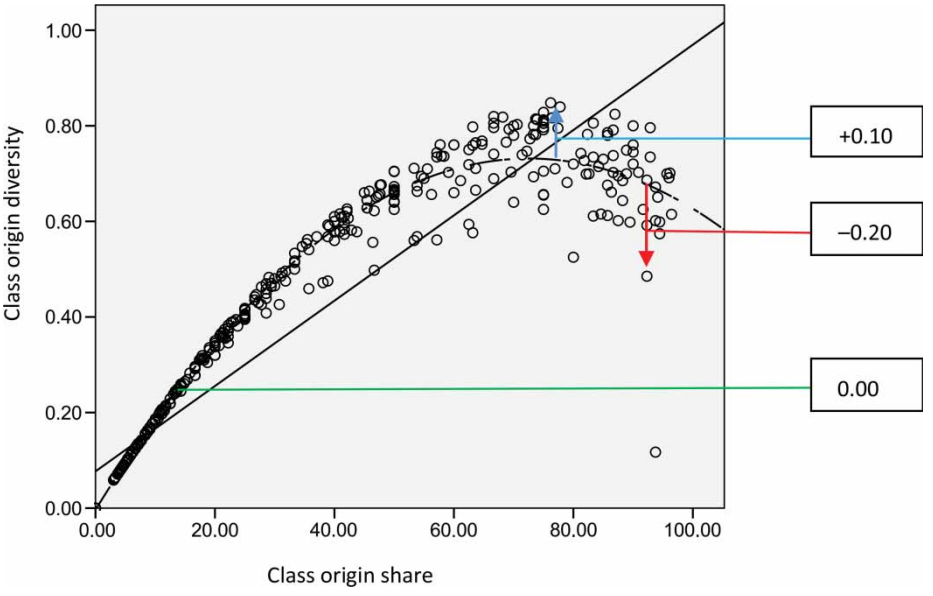


Figure 3.1. Class origin share and origin diversity groups for native students in Grade 8. Source: COOL 2008, own computation.

Appendix 4. Two-level multilevel analyses

Table 4.1. The effects of the origin composition on reading comprehension of students with a native background.

	Grade 2			Grade 5			Grade 8		
	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4
Constant	76.3**	76.3**	76.3**	30.9**	30.9**	30.9**	60.6**	60.6**	60.5**
	(0.4)	(0.4)	(0.4)	(0.5)	(0.5)	(0.5)	(0.8)	(0.8)	(0.8)
Grade level									
Proportion migrants of schoolgrade	–5.7**	–5.5**	–5.4**	–3.9**	–3.7**	–3.5**	–3.2**	–3.0**	–3.0**
	(0.6)	(0.6)	(0.6)	(0.9)	(0.9)	(0.9)	(1.3)	(1.3)	(1.3)
Residualized origin diversity of school grade		–14.9**			–12.0			–18.3	
		(4.8)			(8.0)			(10.4)	
Proportion parents with tertiary education of school grade	–0.6	–0.6	–0.4	2.0*	2.0*	2.4*	4.5**	4.6**	4.9**
	(0.6)	(0.6)	(0.6)	(0.9)	(0.9)	(0.9)	(1.7)	(1.7)	(1.7)
Residualized number of origin groups			–0.3**			–0.4**			–0.5
			(0.1)			(0.2)			(0.3)
Individual characteristics									
Parental education missing	–3.1**	–3.1**	–3.0**	–4.8**	–4.8**	–4.8**	–5.6**	–5.6**	–5.6**
	(1.1)	(1.1)	(1.1)	(1.2)	(1.2)	(1.2)	(1.7)	(1.7)	(1.7)
Low parental education	–6.2**	–6.2**	–6.2**	–12.9**	–13.0**	–13.0**	–15.3**	–15.4**	–15.4**
	(0.8)	(0.8)	(0.8)	(1.2)	(1.3)	(1.3)	(1.5)	(1.5)	(1.5)
Low secondary parental education	–4.4**	–4.4**	–4.4**	–9.5**	–9.5**	–9.5**	13.5**	13.5**	13.5**
	(0.3)	(0.3)	(0.3)	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)
Upper secondary parental education	–2.2**	–2.2**	–2.2**	–5.7**	–5.7**	–5.7**	–7.4**	–7.4**	–7.4**
	(0.2)	(0.2)	(0.2)	(0.4)	(0.4)	(0.4)	(0.4)	(0.4)	(0.4)
Tertiary parental education	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Female	1.4**	1.4**	1.4**	2.9**	2.9**	2.9**	3.0**	3.0**	3.0**
	(0.2)	(0.2)	(0.2)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)
Variance									
Individual level	71.2**	71.3**	71.3**	164.4**	164.4**	164.3**	203.5**	203.4**	203.4**
	(2.1)	(2.1)	(2.1)	(4.5)	(4.5)	(4.5)	(3.4)	(3.4)	(3.4)
School level	18.3**	18.1**	18.1**	11.6**	11.5**	11.5**	22.3**	22.3**	22.1**
	(2.0)	(2.0)	(2.0)	(3.8)	(3.8)	(3.8)	(3.4)	(3.4)	(3.4)
Log likelihood	65115.1	65105.7	65104.8	64920.2	64918.0	64913.4	62141.0	62138.9	62138.8
N students	8895			8108			7568		
N schools	499			503			482		

Notes: Standard errors between brackets. Significance: $^{**}p < 0.01$; $^{*}p < 0.05$.

Source: COOL 2008, own computation.

Table 4.2. The effects of the origin composition on reading comprehension of students with a migrant background.

	Grade 2			Grade 5			Grade 8		
	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4
Constant	67.9**	67.8**	68.1**	22.3**	22.3**	22.3**	52.1**	52.1**	52.1**
	(0.9)	(0.9)	(0.9)	(1.2)	(1.2)	(1.2)	(1.7)	(1.7)	(1.7)
Grade level									
Proportion migrants	–0.5	–0.3	–0.3	–0.8	–0.7	–0.5	–1.7	–1.6	–1.7
of class	(0.7)	(0.7)	(0.7)	(1.0)	(1.0)	(1.0)	(1.6)	(1.5)	(1.5)
Residualized origin diversity of school		–9.1**			3.0			–16.0**	
grade		(2.4)			(2.5)			(6.0)	
Proportion parents with tertiary education	2.8*	3.2**	3.6**	5.6**	5.7**	6.0**	6.3*	6.4*	6.7*
of school grade	(1.2)	(1.2)	(1.2)	(1.7)	(1.7)	(1.7)	(2.8)	(2.8)	(2.8)
Residualized number			–0.6**			–0.3*			–0.4
of origin groups			(0.1)			(0.1)			(0.3)
Individual characteristics									
Parental education missing	–2.3	–2.2	–2.4	–4.9**	–4.9**	–4.9**	–7.3**	–7.3**	–7.4**
	(1.6)	(1.6)	(1.6)	(1.7)	(1.7)	(1.7)	(2.1)	(2.1)	(2.1)
Low parental education	–3.4**	–3.4**	–3.7**	–4.8**	–4.8**	–4.8**	–6.9**	–6.9**	–6.8**
	(0.5)	(0.5)	(0.5)	(0.8)	(0.8)	(0.8)	(0.9)	(0.9)	(0.9)
Low secondary parental education	–2.3	–2.3	–2.3	–4.6**	–4.6**	–4.6**	–7.0**	–7.0**	–7.0**
	(0.5)	(0.5)	(0.5)	(0.8)	(0.8)	(0.8)	(0.9)	(0.9)	(0.9)
Upper secondary parental education	–1.1*	–1.1*	–1.0*	–2.4**	–2.5**	–2.4**	–3.1**	–3.1**	–3.1**
	(0.5)	(0.5)	(0.5)	(0.7)	(0.7)	(0.7)	(0.9)	(0.9)	(0.9)
Tertiary parental education	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Female	0.7*	0.7*	0.7*	2.3**	2.3**	2.3**	2.6**	2.6**	2.6**
	(0.3)	(0.3)	(0.3)	(0.4)	(0.4)	(0.4)	(0.5)	(0.5)	(0.5)
First-generation migrant	–1.8**	–1.7**	–1.7**	–2.0*	–2.0*	–2.0*	–1.4	–1.4	–1.4
	(0.7)	(0.7)	(0.7)	(0.8)	(0.8)	(0.8)	(0.8)	(0.8)	(0.8)
Turkish origin	–1.9**	–2.1**	–2.1**	–2.0*	–1.9*	–2.0*	–2.0*	–2.1*	–2.0*
	(0.5)	(0.5)	(0.5)	(0.8)	(0.8)	(0.8)	(0.9	(0.9	(0.9
Moroccan origin	–0.2	–0.3	–0.4	–0.2	–0.2	–0.3	0.7	0.6	0.6
	(0.5)	(0.5)	(0.5)	(0.8)	(0.8)	(0.8)	(1.0)	(0.9)	(0.9)
Western origin	4.3**	4.3**	4.2**	2.6*	2.6*	2.5*	2.8*	2.8*	2.7*
	(0.8)	(0.8)	(0.8)	(1.1)	(1.1)	(1.1)	(1.4)	(1.4)	(1.4)

(*Continued*)

Table 4.2. Continued.

	Grade 2			Grade 5			Grade 8		
	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4
Eastern-European origin	–1.5	–1.5	–1.4	0.2	0.3	0.3	3.2*	3.2*	3.2*
	(0.8)	(0.8)	(0.8)	(1.2)	(1.2)	(1.2)	(1.4)	(1.4)	(1.4)
Chinese origin	–1.4	–1.3	–1.1	4.8*	4.8*	4.9*	4.8*	4.8*	4.9*
	(1.4)	(1.4)	(1.4)	(1.9)	(1.9)	(1.9)	(2.2)	(2.2)	(2.2)
Iraqi origin	–1.5	–1.4	–1.3	–1.5	–1.5	–1.5	0.6	0.6	0.7
	(1.0)	(1.0)	(1.0)	(1.7)	(1.7)	(1.7)	(1.8)	(1.8)	(1.8)
Afghan origin	0.0	0.0	–0.1	1.9	1.9	2.0	6.1**	6.1**	6.1**
	(1.3)	(1.3)	(1.3)	(2.1)	(2.1)	(2.1)	(2.0)	(2.0)	(2.0)
Somali origin	0.9	1.1	1.0	2.4	2.5	2.4	–0.4	–0.5	–0.5
	(1.0)	(1.0)	(1.0)	(1.7)	(1.7)	(1.7)	(2.3)	(2.3)	(2.3)
Other country origin	–0.1	–0.1	–0.1	2.6	2.6	2.5	3.3**	3.2**	3.3**
	(0.6)	(0.6)	(0.6)	(0.8)	(0.8)	(0.8)	(0.9)	(0.9)	(1.0)
Former colony origin	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Variance									
Individual level	58.8**	58.6**	58.6**	115.4**	115.4**	115.8**	176.9**	179.7**	176.6**
	(2.6)	(2.6)	(2.6)	(5.4)	(5.4)	(5.4)	(5.1)	(5.1)	(5.1)
School level	20.1**	19.9**	19.3**	17.1**	17.0**	16.5**	21.8**	20.4**	21.4**
	(2.6)	(2.6)	(2.6)	(4.8)	(4.8)	(4.8)	(3.7)	(3.6)	(3.7)
Log likelihood	24922.3	24907.4	24883.2	21404.3	21402.8	21399.9	22751.4	22744.4	22749.2
N students	3469			2773			2808		
N schools	374			378			369		

Notes: Standard errors between brackets. Significance: $^{**}p < 0.01$; $^{*}p < 0.05$.
Source: COOL 2008, own computation.