

CSAE Working Paper WPS/2018-16

Ethnic diversity in Peruvian schools: Disentangling

peer and class composition effects

Nicolás Pazos Navarro

University of Oxford

npazosn@gmail.com

Abstract

Academics have thought of many ways to reduce the gaps in academic achievement for ethnic minorities, one of which is through manipulating the classroom composition. Many studies show that there are positive spill-overs or peer effects of sharing space with better performing students (Sacerdote, 2001). However, other studies find that students from ethnic minorities tend to be negatively affected by ethnically diverse classes, either because of racial stereotypes or segregation (Steele and Aronson, 1995). This study applies Lee's (2007) linear-in-means model with fixed effects to a sample of approximately 140 thousand students from a national education census in Peru to disentangle endogenous peer effects from contextual peer effects, while solving for self-selection and the reflection problem. The study focuses on the effects of ethnic diversity on Peruvian indigenous students. Results suggest that the effect of class composition on ethnically diverse classes is driven by two distinct and opposing channels, one related to abilities spill-overs and one to stereotypes and segregation. Endogenous peer effects are negative, possibly due to incentives faced by teachers and school directors to focus on better students and neglect the poorer performing ones, thus enlarging the gap between them. In terms of ethnic composition, results are heterogeneous: indigenous students benefit from a large proportion of similar peers, but the effect is the opposite for non-indigenous students. The study concludes that the current composition of the average Peruvian class is suboptimal, and a larger proportion of indigenous students could help reduce the current gaps in achievement due to ethnicity.

1. Introduction

Educational achievement has been on the rise in Peru during the last decade, mostly because of the country's favourable macroeconomic climate and major educational reforms. However, students' achievement varies tremendously according to socioeconomic status and ethnic background, perpetuating the already high levels of inequality. By 2016, only 2.3% of indigenous students attending second grade of secondary school had achieved a satisfactory level for mathematics in the ECE evaluation (a national census for measuring quality of education in schools), compared to 11.4% when looking at the whole population. At the same time, only 1.5% and 2.6% of indigenous students achieved a satisfactory level for reading and history, geography and economy, respectively, compared to 14% and 15% of the whole population.

The indigenous population has been historically marginalized, typically living in rural areas, or having migrated and living on the outskirts of big cities¹; and facing high incidence of poverty, malnourishment and lack of access to basic services². These characteristics are important drivers of their poor academic performance, but evidence shows individual characteristics are not the only determinants of education achievement. Studies like the ones from Sacerdote (2001) and Carrell et al (2009) show positive peer effects of sharing space with better performing students, and suggest mixing good performing with bad performing students could benefit the latter. This applies to ethnic minorities, who usually perform worse than their peers. However, peer effects studies are inconclusive: When positive, peer effects are moderate, small or insignificant, and sometimes even negative (Foster, 2006; Carrell et al., 2013). On the other hand, psychological related literature (Steele and Aronson, 1995; Inzlicht and Ben-Zeev, 2000) shows that students with ethnic backgrounds tend to be negatively affected by diverse spaces, either because of racial stereotypes or segregation.

Given the indigenous students' characteristics, there are at least two components driving the effect of ethnic composition over academic achievement. First, students with an indigenous background may be positively (or negatively) affected by sharing class with students with better academic abilities. This effect is related to academic background and abilities of the students, and not directly with ethnicity. Second, racial stereotypes and discrimination in

¹ Migration from rural to urban cities is common: many indigenous people now live in cities, usually settling in the outskirts, where they get some improvement in access to services but also face segregation and poverty.

² Data from the 2016 household survey (ENAH0) shows income in Spanish households is 50% larger than in indigenous households. By this year, 30% of indigenous population was still living under the poverty line.

classes may affect the results of indigenous students, reducing their motivation and causing stress. Using data of Peruvian schools, this study analyses how the proportion of indigenous children in classrooms affect their academic outcomes, disentangling the multiple peer effects involved.

Identifying class composition effects is a complicated task, since three different effects must be identified: *Endogenous peer effects*, where the results of the individual change through a change in the peers' results; the *contextual peer effects*, where the results of the individual change given the characteristics of the peers; and the *correlated effects*, where members of a group tend to behave similarly given their characteristics.

As detailed by Manski (1993), the existence of these three different effects originates three potential problems. The first problem is self-selection, which makes it difficult to separate correlated effects from peer effects. In my study, we might be concerned about self-selection into schools and class-rooms. Indigenous students usually live in poorer areas, and attend poorer schools. Also, children might be separated into classes within the same school according to certain characteristics, either because of specific school policies, or because more influential parents may lobby to get their children in classes with better teachers. Thus, indigenous students may be selected into classes with worse peers and teachers, making it common to find that classes with a larger proportion of indigenous students perform worse. This, however, is not an effect of class composition, but a consequence of how classes are formed. The second problem is the reflection problem, which makes it difficult to differentiate between contextual and endogenous peer effects³. Addressing this problem is of particular relevance for this study, since its main goal is to differentiate between endogenous (ability) and contextual (ethnicity) peer effects, which are correlated. A third problem is the identification of the relevant groups. The analysis assumes a priori that classes are the relevant groups.

The main objective of these paper is to test the existence and analyse the channels through which class composition in terms of ethnicity affects results, with a focus in how to improve results for indigenous students. To do so, it is necessary to solve for self-selection and the reflection problem, and to isolate the endogenous peer effects from other effects produced by ethnic diversity. I use a method developed by Lee (2007) on data from the ECE census -

³ The contextual peer effects are all other characteristics that may influence the peers' results that are not the endogenous effect (e.g. ethnicity, socioeconomic background).

an evaluation done to all students in second grade of secondary school – to solve the self-selection and reflection problems, and to separately estimate endogenous and contextual peer effects. This method exploits variation in group size and uses class fixed effect to identify the aforementioned effects. Additionally, I add an interaction effect between class composition and ethnicity to allow for heterogeneities. This is the first time heterogeneous effects are included in the derivation of Lee’s model, making it a contribution to the model and to the peer effects literature. I report results for mathematics, reading, and history, geography and economy for a sample of approximately 140 thousand students in ethnic diverse classes (at least one indigenous and one non indigenous student) of second secondary grade.

After solving the reflection problem, results confirm the existence of two opposing channels: one endogenous related to abilities, and one exogenous related to ethnicity and other characteristics. The estimates show negative endogenous peer effects, possibly due to the incentives faced by the teachers and school directives to focus on better students and neglect bad ones, thus enlarging the gap between them, or as a consequence of within-classroom group formation. Moreover, results of ethnic composition are heterogeneous: the proportion of indigenous students in class has a small positive effect on indigenous students, but larger negative effects on the others. Finally, the study concludes that the current composition of the average Peruvian class is suboptimal, and a redistribution to achieve larger proportion of indigenous students could help reduce the current gaps in achievement.

To my knowledge, this is the first study to disentangle the channels through which ethnic diversity in class rooms affects academic outcomes. As will be explained in the next section, previous studies trying to identify the effects of ethnic diversity fail to disentangle the channels correctly, possibly leading to biased conclusions about the effects of class composition. Considering the different characteristics of ethnic groups and their effect on their peers is fundamental to understand the effects of class composition in its full dimension. This is also the first quantitative study to address the problem of indigenous people in ethnic diverse classes in Peru. At the same time, it is the first time Lee’s model has been used in the context of a developing country.

The paper is divided as follows. Section 2 briefly discusses the previous literature and the conceptual framework; Section 3 explains the methodology; Section 4 presents the Data; Section 5 shows the empirical results; and Section 6 interprets the results. Section 6 is divided in 2 subsections: Subsection 6.1 explains the possible mechanisms and 6.2 discusses the

optimal class composition. Section 7 concludes. The simulations for the CML estimator, and a short analysis of the presence of ethnic segregation in classes are shown in the Appendix.

2. Literature Review and Conceptual Framework

This paper is based on the peer effects literature and the racial stereotypes literature. The study proposes a link between both literatures, where contextual effects like the ethnic composition of the class and endogenous peer effects related to ability are both channels through which ethnic diversity in class composition affects results.

The first literature is usually concerned about finding the endogenous peer effects, based on the idea that sharing a classroom with better achieving students can lead to a multiplicative positive (or negative) effect on each individual student. Measuring peer effects, however, is a complicated task: the researcher needs to specify the reference groups, then solve for self-selection to separate peer effects from correlated effects, and finally address the reflection problem (Manski, 1993) to separate endogenous from contextual peer effects.

Though many studies use data from students' social networks (friends) to address the problem of relevant groups, it is a common practice to analyse peer effects at a class level (Ammermueller and Pischke, 2009) due to data availability. To solve the problem of self-selection, some studies take advantage of cases where students are randomly selected into groups. Sacerdote (2001) exploits the fact that new students at Dartmouth College are randomly assigned to their roommates, and finds that sharing room with a student with good academic results has a positive impact in academic results at the end of the year. Kang (2007) exploits quasi-randomization of students in South Korean's middle-schools, finding a positive correlation between students' results and their peers' mean achievement. Other studies have focused on estimating peer effects with observational data, using fixed effects at a group level to solve the self-selection problem. To avoid the peer effects from being absorbed in the fixed effects, it is usually necessary to define fixed effects at a larger level than the reference group. In this way, Ammermueller and Pischke (2009) used fixed effects at a school level to find peer effects at a class level. This, however, requires variation between classes to be exogenously determined, so selection into classes should be random.

Although providing interesting results, all of these studies fail to truly address the reflection problem, since they are unable to disentangle endogenous from contextual peer effects. Instead, they are implicitly assuming the inexistence of contextual effects, or estimating a

reduced form model where the parameter includes both effects. Lee (2007) proposes a model that finally solves this problem. In his model, peer groups are defined for each individual as the group excluding the individual itself. Thus, each member of the group will have his own reference group within the group, permitting the separation contextual and endogenous peer effects while using a within estimation. Lee's model will be discussed in more detail in the methodology section. On this matter, Boucher et al (2014) use Lee's model on secondary school students in Canada, finding positive endogenous peer effects. No study using this methodology has been conducted in low- or middle-income countries.

On the other hand, the literature about racial stereotypes and other kinds of discrimination in ethnic diverse classrooms is related to the contextual peer effects channel. Steele and Aronson (1995) and Steele (1997) find that racial stereotypes explain an important part of low academic performance in black students in the USA, even when controlling for socioeconomic background. Further studies showed that the negative effects on stereotyped minorities are higher when the presence of other ethnic groups is large. Inzlicht and Ben-Zeev (2000) found that math scores for a group of girls diminished when male peers were introduced to the classroom, increasing the gap between boys and girls. Inversely, a study by Marx and Roman (2002) showed that girls improved their math test's results when visited by a female examiner. Both of these studies show how the presence and representativeness of different groups in the classroom can influence the outcomes by reinforcing social stereotypes and causing stress. However, stereotypes not only generate bad results because of stress: They can also lead to fear, segregation, harassment and bullying. Gray-Little and Hafdahl (2000) find that classes with more combinations of different distinct ethnic groups may have negative effects on self-esteem. Agirdag et al (2011) find that students from ethnic minorities are less prone to be victims of segregation when surrounded by more students from the same ethnic group. Buhs et al (2006) show segregation can lead to negative effects on academic results. As shown, ethnic diversity in classes can lead to worse academic outcomes for minorities through many mechanisms related to stress and self-esteem.

Many studies document the presence of these stereotypes and other forms of discrimination in Peruvian schools (Valdivia, 2003; Valdivia et al., 2007). A study conducted by Glewwe et al (2014) finds that indigenous students learn less in schools than their peers, even when attending the same education centre. Kudo (2004) and Rodriguez Lozano (2012) find it is more probable for these students to work and study, repeat grades and abandon school. To

my knowledge, there have not being any studies about the effects of racial stereotypes on academic achievement conducted to this specific population.

Following Manski's reasoning, the existence of multiple peer effects may explain why some studies end up contradicting each other. For example, Angrist and Lang (2004) used evidence from an integration program in the United States that increased the number of students from minority groups in previously less diverse schools (thus increasing the number of low-performing students). They do not find any negative peer effects on white students' results, but do find some negative effects on black female students. These results suggest that either there are negative contextual peer effects (driven by stereotypes), endogenous peer effects are negative, or both. On the existence of negative peer effects, Foster (2006) finds small negative peer effects measured as SAT score on men's GPAs. Carrell et al (2009) use exogenous variation to peer groups and find that an increase in the peer average score for verbal SAT increases each individual's results, suggesting positive peer effects. However, when in a paper by Carrell et al. (2013) they use these same results to endogenously assign peer groups to maximize the performance of lower ability students, they find negative peer effects.

The main objective of this study is to understand and disentangle the endogenous peer effects from the contextual effects in terms of ethnicity. No studies have been conducted to address the specific question of ethnic diverse classes in Peru and its effect on indigenous students. A study by Cueto et al. (2016) finds positive effects of classroom composition according to parents' education level over academic achievement and socioemotional abilities in Peru, but they do not disentangle endogenous from contextual peer effects. The following section will explain the methodology used to solve this issue.

3. Methodology

The methodological basis for this work builds on Lee (2007)⁴. Particularly, this study uses Lee's linear-in-means model with fixed effects, estimated with a conditional maximum likelihood. The key assumptions, main equations and an explanation of the econometric methods are presented in the following paragraphs. Notation has been changed from Lee's paper, following Boucher et al (2014) approach to the model, to facilitate its reading.

⁴ Bramoullé et al. (2009) and Boucher et al. (2014) work over Lee's model were also considered in the development of this study.

Some key assumptions are needed for this model to work. First, it is assumed that groups are defined by classes, and that students interact only with members of their class. Even if this is not the first study to define groups a priori (Boucher et al., 2014), it is still a strong assumption. It is possible for students to be interacting at school or district level, and it is also possible that sub groups are endogenously formed within classes. Given the nature of the data, it is not possible to identify exactly which students are interacting with each other.

Second, Lee's model assumes that each student's peer group is different from the whole group (the class) by excluding the individual herself. This assumption is key, since it is necessary to allow the use of class-level fixed effects, and to solve the reflection problem. The third assumption is that the linear-in-means model is sufficient to evaluate each student's resulting score. The linear-in-means model proposed looks as follows:

$$y_{r,i} = \alpha_r + \beta \frac{\sum_{j \in P_i} y_{r,j}}{m_r - 1} + \boldsymbol{\gamma} \mathbf{x}_{r,i} + \boldsymbol{\delta} \frac{\sum_{j \in P_i} \mathbf{x}_{r,j}}{m_r - 1} + \rho z_{r,i} + \varepsilon_{r,i} \quad (1)$$

where $y_{r,i}$ is test score for student i in class r , which depends on the mean of her classmates' test score $\frac{\sum_{j \in P_i} y_{r,j}}{m_r - 1}$ (not including student i), some individual characteristics described by vector $\mathbf{x}_{r,i}$, the mean of his peers characteristics $\frac{\sum_{j \in P_i} \mathbf{x}_{r,j}}{m_r - 1}$, and an interaction term $z_{r,i} = x_{1r,i} \frac{\sum_{j \in P_i} x_{1r,j}}{m_r - 1}$, where x_1 represents the first input of the $\mathbf{x}$ vector: ethnicity. Parameter α_r captures all group invariant unobserved variables, and $\varepsilon_{r,i}$ is the error term. The coefficient β captures the endogenous peer effect, while the vector of coefficients $\boldsymbol{\delta}$ captures the exogenous peer effects, including the effect of the class composition in terms of ethnicity. The vector of coefficients $\boldsymbol{\gamma}$ captures individual effects, and ρ captures the effect of the interaction between ethnicity and the proportion of indigenous students in class. This interaction term was not included in Lee's original model or in any other paper following his strategy, but it is a contribution of this study.

To eliminate parameter α_r , let us use a within transformation. Let us assume $E(\varepsilon_{r,i} | \mathbf{X}_r, m_r \alpha_r) = 0$. The within reduced form equation is:

$$y_{r,i} - \bar{y}_r = \frac{\gamma - \frac{\delta}{m_r - 1}}{1 + \frac{\beta}{m_r - 1}} (\mathbf{x}_r - \bar{\mathbf{x}}_r) + \frac{\rho}{1 + \frac{\beta}{m_r - 1}} (z_r - \bar{z}_r) + \frac{1}{1 + \frac{\beta}{m_r - 1}} (\varepsilon_{r,i} - \bar{\varepsilon}_r) \quad (2)$$

All means of this model are computed for all students in the group (that includes student i). It is worth noticing that the effect of class size is controlled by the fixed effect. Assuming

that the individual is excluded from his own peer groups allows distinguishing peer effects from correlated effects, and endogenous from contextual peer effects in the structural equation. Then, I will need a method to recover the structural parameters.

The model will be estimated using a conditional maximum likelihood estimator (CML). As explained by Lee (2007) a within OLS would behave properly only when the sizes of groups are much larger than the total number of groups, a property that is violated in our data. Also, when class sizes are too big, the CML would converge to the OLS, leading to biased results.

Before deriving the log-likelihood function, we correct for missing values, as done in Davezies et al. (2009) and implemented in Boucher et al. (2014). Even when all students were supposed to take the exam, some variables are missing. However, most of the missing variables are found in socioeconomic and ethnic background variables, and not in the exams' results. There is no reason to assume this represents any systematic bias. The total number of missing observations represents less than 5% of the total sample. Before doing the correction specified in Davezies et al. (2009), classes with more than two missing observations⁵ will be excluded to prevent possible biases. The log-likelihood function, after correcting for missing values, and assuming a normal distribution for the error, looks as follows:

$$\ln L = c + \sum_{r=1}^R (n_r - 1) \ln(m_r - 1 + \beta) - \frac{N - R}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{r=1}^R \left(\frac{m_r - 1 + \beta}{m_r - 1} \mathbf{y}_r^* - \mathbf{X}_r^* \frac{(m_r - 1)\gamma - \delta}{m_r - 1} - \rho \mathbf{z}_r^* \right)' \left(\frac{m_r - 1 + \beta}{m_r - 1} \mathbf{y}_r^* - \mathbf{X}_r^* \frac{(m_r - 1)\gamma - \delta}{m_r - 1} - \rho \mathbf{z}_r^* \right) \quad (3)$$

Where $\mathbf{a}_r^* = (\mathbf{a}_r - \bar{\mathbf{a}}_r)$, c is a constant, m_r is the total number of students in class r , n_r represents the number of non-missing peers in group r ⁶, N is the total number of non-missing observations, R is the total number of classes, and σ^2 is the error variance. Variables $\mathbf{y}_r^*$, $\mathbf{X}_r^*$, $\mathbf{z}_r^*$ are the matrix equivalents to the variables presented in equation 2.

The parameters are recovered using the Matlab's `fmincon` function and a programme written specifically for this study. There was no online version of Lee's estimator, so the writing of the program was done entirely by the author of this paper. To be sure the program was working properly and recovering the right estimators, a number of simulations were

⁵ A series of simulations were performed to understand how missing observations could influence the results.

⁶ So it follows that $\forall r, n_r \leq m_r$.

performed. The simulations and all the information about how they were implemented can be found in the Appendix.

4. Data

The data used comes from the Peruvian census of educational attainment⁷ (ECE, for its name in Spanish). Each year, Peruvian students from 2nd grade of primary and 2nd grade of secondary take a test that measures their skills in mathematics and language, which is evaluated in Spanish. In year 2016, a test for history, geography and economy was added. The scores are analysed by the ministry of education following the Rasch model, with mean 500. The suggested way of interpreting the scores is described in Table I. The exam is compulsory, and all students are required to take it. There is no formal punishment for students who perform badly in the exam, but the results at the school level are then used by the Ministry of Education as a criteria for allocating resources (Budgeting for Results). Besides the exam, students answer a questionnaire about their gender, socioeconomic and family backgrounds, and other variables. This questionnaire is self-assessed, and does not has any impact on the scores. The census also recovers information about the class distribution, and each student can be linked to a specific class, grade and school.

Table I: Interpreting Scores

	Below initial point	Initial point	In process	Satisfactory
Math score	<520	[520,595]	[596,648]	648<
Reading score	<505	[505,580]	[581,640]	640<
History score	<426	[426,499]	[500,606]	606<

Source: MINEDU UMC (Ministry of Education, Peru)

I only used data for second secondary school in the analysis. Primary schools in Peru are more common and more dispersed than secondary schools, so students can attend primary school in their own localities, but in many cases are forced to leave their communities to attend secondary school in larger cities or towns. Also, Intercultural Bilingual Education (EIB) is available only for primary schools, and not for secondary education. Thus, students coming from an indigenous background can do their primary education studies at their own communities and in their own languages, and only be confronted with students from different backgrounds once they reach secondary education.

The total population of students attending second grade of secondary school in Peru is around half a million. This study focuses on multiculturalism in schools, so only data of

⁷ The data set is available online upon request.

classes with at least one indigenous student and one non-indigenous student is analysed. Indigenous background is measured using a question about their parents' main language: A child is considered as having an "ethnic background" if at least one of her parents spoke with her in an indigenous language⁸. Also, the final sample includes only classes for which a maximum of two observations are missing, and with at least more than 5 students (thus, no class has more than 50% missing observations). Assuming missing observations are randomly generated, it would be more probable to encounter missing observations in larger classes than in smaller classes. Given the size of the sample, getting rid of missing information should not represent a problem in terms of statistical power. The final sample has 141,578 students, distributed across 4,766 schools and 6,890 classes.

Table II: Summary Statistics

	Total Sample		Non-indigenous students		Indigenous students		Diff
	Mean	SD	Mean	SD	Mean	SD	
Maths Score	544.72	77.05	555.96	76.45	506.82	66.22	-49.14***
Language score	550.25	66.43	562.07	65.21	510.39	53.81	-51.68***
History score	482.11	96.37	496.15	95.32	434.75	84.04	-61.4***
Male	0.51	0.50	0.51	0.50	0.50	0.50	-0.01**
Socioeconomic status index	-0.38	0.93	-0.18	0.89	-1.05	0.72	-0.88***
Socioeconomic status (high)	0.18	0.38	0.22	0.41	0.03	0.16	-0.19***
Socioeconomic status (med)	0.34	0.47	0.39	0.49	0.15	0.36	-0.24***
Harassment dummy	0.26	0.44	0.24	0.43	0.32	0.47	0.08***
Repeated a year	0.23	0.42	0.21	0.40	0.33	0.47	0.13***
At least one parent finished secondary school	0.60	0.49	0.67	0.47	0.36	0.48	-0.31***
Rural	0.17	0.37	0.11	0.32	0.34	0.47	0.23***
Class Size	24.68						
School Size	103.78						
Number of Observations	141,578		109,199		32,379		
Proportion of total sample	1.00		0.77		0.23		
Number of classes	6,890						
Number of schools	4,766						

Note: .01 - ***; .05 - **; .1 - *

Table II shows the descriptive statistics of the final data. As we can see, on average students perform better in language than in maths or history. The low score in history may be caused by the lack of attention that the Peruvian system gave to this topic, in favour of a much more intensive focus on maths and reading skills. Also, 2016 was the first year in which history was evaluated as part of the census. For all three results, students with an indigenous

⁸ Just 0.1% of the children spoke only foreign languages (typically English) with their parents, and were counted as not having an ethnic background (same with Spanish speakers).

background perform significantly worse than their peers. History and language are the courses where larger differences between students from different ethnic backgrounds are found, while differences in math are still important. This might be related to differences in their ability to speak Spanish. Indigenous students are also significantly poorer, and on average they go more to schools located in rural areas than non-indigenous students. Students with an indigenous background are also typically more likely to be victims of harassment in schools⁹, being harassed significantly more than their peers by 8%. This could serve as weak evidence for the social segregation described in the literature, and will be discussed with more detail in the Appendix.

Class rooms in the sample have a mean of approximately 25 students per class, and a standard deviation of 8. As explained by Boucher et al. (2014), the variation in class sizes would be useful at the moment of estimating the parameters. Average school size in the data is of 104 students¹⁰. Only 23% of the sample falls into the category of indigenous, which is not surprising given their minority status.

5. Empirical Results

Table III reports the results of the conditional maximum likelihood model for math, reading and history scores. The estimation was done following equation (3). The conditional maximum likelihood recovers the structural parameters. The contextual peer effects coefficients are measured with respect to the percentage of peers characteristics. In other words, the coefficients capture the effect of a 100% change in the variable, (going from no peers sharing the specific characteristic to all peers sharing them). For this reason, it is not surprising to find effects with larger coefficients than the individual effects as the change is large. Results in columns 1, 4, and 7 include only endogenous peer effects and contextual and individual effects in terms of ethnicity. Columns 2, 5, and 8 add more controls at an individual level and more contextual peer effects. Finally, columns 3, 6, and 9 include heterogeneous effects by including the interaction between individual ethnicity and class composition in terms of ethnicity.

⁹ The harassment index was created using seven items included in the questionnaire. Those items include verbal and physical harassment, both performed by teachers and classmates.

¹⁰ When looking at the whole data set, it is possible to notice that students with an indigenous background tend to attend to significantly smaller schools and classes, with an average size of 65 and 21 students, respectively. When reducing our sample to only mixed classes, the difference becomes 0 by construction.

Table III: CML estimation on Scores

	Math scores			Reading scores			History scores		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Endogenous peer effect	7.942***	-1.960***	-1.997***	6.455***	-1.999***	-1.998***	5.977***	-1.997***	-1.999***
(peer avg. score in class)	(0.056)	(0.000)	(0.000)	(0.018)	(0.000)	(0.000)	(0.070)	(0.000)	(0.000)
Contextual peer effects									
Indigenous	96.275***	-49.246***	-105.171***	53.596***	-54.324***	-62.820***	55.641***	-52.210***	-92.515***
(% in class)	(13.979)	(11.941)	(6.991)	(6.255)	(10.356)	(7.841)	(0.891)	(13.837)	(15.341)
Male		38.377***	23.684***		-32.276***	-53.297***		-49.331***	-30.305***
(% in class)		(10.190)	(7.857)		(5.44)	(8.021)		(14.123)	(7.549)
Socioeconomic Status: High		-121.575***	-201.973***		-152.945***	-238.804***		129.425***	34.613***
(% in class)		(21.764)	(20.297)		(14.176)	(19.282)		(15.029)	(10.221)
Socioeconomic Status: Medium		-19.948	-30.14**		2.86	-99.936***		95.254***	69.69***
(% in class)		(15.638)	(15.201)		(10.537)	(10.162)		(14.518)	(13.309)
At least one parent finished		77.508***	57.693***		89.446***	132.11***		60.414***	28.801***
secondary school		(12.034)	(12.933)		(9.986)	(7.845)		(15.577)	(9.883)
(% in class)									
Individual effects									
Indigenous	-24.871***	-19.115***	-26.463***	-26.596***	-20.068***	-25.597***	-32.862***	-24.462***	-33.205***
(yes=1)	(0.993)	(0.714)	(0.543)	(1.318)	(0.584)	(0.492)	(7.771)	(0.789)	(0.929)
Male		14.561***	13.597***		-2.934***	-4.126***		1.32	2.100***
(yes=1)		(0.520)	(0.395)		(0.350)	(0.450)		(0.838)	(0.334)
Socioeconomic Status		3.463***	-0.11		4.77***	0.579		14.105***	9.823***
(High=1)		(0.973)	(0.887)		(0.669)	(0.834)		(0.818)	(0.557)
Socioeconomic Status		4.308***	3.913***		6.096***	1.061**		10.044***	8.650***
(Medium=1)		(0.753)	(0.747)		(0.481)	(0.500)		(0.759)	(0.734)
At least one parent finished		9.403***	8.466***		11.497***	13.672***		11.149***	9.146***
secondary school (yes=1)		(0.665)	(0.716)		(0.585)	(0.415)		(0.905)	(0.609)
Heterogeneous effect									
Indigenous individual effects *			14.253***			17.043***			13.609***
Indigenous contextual peer			(1.057)			(0.900)			(1.998)
effects									

Note: .01 - ***; .05 - **; .1 - *; Standard errors in brackets. All coefficients have been rounded up to a maximum of three decimals.

For the first group of results (columns 1, 4, and 7), we see that being indigenous has a significant negative effect on the results of the three scores, but having indigenous peers has a positive and significant effect. At the same time, endogenous peer effects are significant and positive, which is consistent with most of the literature. As stated in Section 2, most of the literature usually fails to account for the reflection problem. This first model is only partially addressing the reflection problem, since it is only disentangling the endogenous peer effects from the contextual peer effects of class composition in terms of ethnicity. If there is more than one contextual fixed effect correlated with the endogenous peer effects, then the reflection problem would still be playing a part, and our endogenous parameter would be incorporating other effects. The next two models control for more variables to solve any other potential correlation.

The second group of results (columns 2, 5, and 8) include more controls in terms of individual effects and contextual peer effects. All individual effects are consistent with what we would expect: Students coming from a richer, more educated family achieve better; and male children do better in maths but worse in reading tests. However, both the endogenous peer effect and the contextual peer effects of having indigenous students in class become significantly negative for all three models. This change of sign is due to correlations between these variables and the newly added contextual variables. The existence of this correlation is what we call the reflection problem, and the change of sign after solving for it confirms the importance of properly disentangling the effects.

Sharing a class with students with a high socioeconomic status has a significantly negative effect on results for maths and reading, but significantly positive effects on history scores. It would be a mistake to interpret these results as if classes with children from a higher socioeconomic status perform worse in maths and reading exams, since the model solves for self-selection and controls for correlated effects. The within class estimation controls for all effects related to school characteristics and teacher's quality, but it does not control for particular behaviours and unequal teacher-student interactions. Consistent with the results found in Cueto et al (2016) for Peru, individuals sharing class with children of better educated parents perform better in maths and reading scores. They also perform better in history.

The third group of results (columns 3, 6, and 9) adds an interaction between individual effects of ethnicity and contextual peer effects of ethnicity, allowing us to disentangle the differences in contextual effects for indigenous students. This is an addition to Lee's original model, which did not account for any heterogeneity. Results show that having more indigenous

students in class has a positive effect on other indigenous students, for all three scores, which is consistent with the racial stereotypes literature: Having more indigenous peers in class diminishes the effects of stress and segregation. Adding this variable keeps the endogenous peer effects almost intact. Contextual peer effects of ethnic diversity and correlated variables are adjusted after including the heterogeneous effects, but no other major changes are found.

6. Interpreting the results

6.1 Discussion and possible mechanisms

Even when not controlling for other characteristics, the first set of results of the CML model is useful to corroborate the existence of at least two opposing channels through which ethnic diversity in class affects individual scores. As shown, sharing class with better achieving student's increases the educational outcome. The effects shown are very large, with each additional point in the average score accounting for a full 7 extra points in individual scores. Indigenous students usually perform worse than their peers, so having a larger proportion of them as peers is indirectly correlated with lower endogenous peer effects. On the other hand, having more indigenous students in class also improves scores for everybody via the contextual peer effects channel, although these effects are relatively small compared to the endogenous peer effects. This channel may be related to an increment on indigenous students' performance due to a reduction in racial stereotypes. These results are still partially biased, and do not account for the existence of other channels through which interaction in class may affect individual performance. The existence of other channels (the reflection problem) becomes clear in the next two models.

The second and third model include more contextual peer effects, turning the two peer effects found in the first model negative for all output scores. The reason behind these changes in sign is a strong correlation between the recently added contextual peer effects and both scores and ethnicity. It is not surprising to find that contextual peer effects of socioeconomic status and parent's education level are correlated with the ability and ethnicity of the peers: The more indigenous people in the class, the less the average level of parent's education in the class, and the smaller the proportion of students with a high or medium socioeconomic status. It is also a consequence of solving the reflection problem.

By looking at the results, it appears that the main reason why indigenous peers had a positive effect on math and reading scores is because of their low socioeconomic status. One reason

why high socioeconomic status could have negative effects might be because of segregation or stigmatization of poorer students: The greater the proportion of rich students in class, the more stigmatized poor students are. However, positive effects of high socioeconomic status peers on history scores suggest that these effects might be driven by a different channel. The channel proposed has to do with dynamics between students and teachers: Even when controlling for teacher ability, richer parents might still lobby or influence them to pay special attention to their own children, marginalizing other students. Teacher's attention is then divided heterogeneously across students in class according to their socioeconomic status. This effect would be more prevalent in classes with a lower proportion of poor students, since a smaller part of the attention would be assigned to them. The reason why the positive history scores support this explanation is because it probably require less division of attention, making this effects disappear. Also, the history test is new to the 2016 ECE, and dynamics involving teachers and parents might take longer to form in order to have an impact on scores.

The positive heterogeneous effects of ethnic diversity show that the racial stereotypes and segregation channel is still explaining part of the indigenous students' scores. These effects are relatively small when compared to the total contextual peer effects of the proportion of indigenous students. Thus, solving the stigmatization problem through an increase in the proportion of indigenous students will not compensate for other negative effects related to this proportion. Still, it might be worth looking for other policies to eradicate stigmatization. More evidence on the presence of ethnic segregation and the determinants of harassment in classes is shown in Appendix 2.

One of the most surprising results are the negative endogenous peer effects. Apparently, all the positive endogenous peer effects where deeply associated with other contextual variables, like the level of education achieved by parents. This variable might be a good proxy of ability, which explains its positive contextual peer effects, and suggests positive spill-overs of ability. If these spill-overs are being captured by the contextual peer effects, then the endogenous peer effects might be capturing some other dynamics within the class, like heterogeneous interactions between teachers and some students. It might be the case that teachers focus only on their best or average students, hence creating larger differences between them and their lower performing peers. Therefore, the negative endogenous peer effects found could be a result of a behaviour that favours good performing students and neglects the rest.

Another similar interpretation is that classes are not the true reference groups. As found by Carrell et al (2013), subgroups are usually formed within classes. These subgroups usually respond to characteristics of their members, where good performing individuals tend to form groups with similar peers. This is also consistent with microeconomic theory about group formation. Ignoring the real reference group is an important limitation of this study.

6.2 Optimal class composition

As discussed, there are many reasons why giving an interpretation of all the variables involved in the contextual peer effects to try to arrive to the optimal class composition is a difficult task. However, we can still find an approximation of the optimal class composition in terms of ethnicity by making some assumptions and generalizations.

Figure I presents graphs showing endogenous, contextual and total peer effects on individual scores of indigenous students for maths, reading and history tests. The graphs show how the endogenous peer effects and all the contextual peer effects interact according to the proportion of indigenous students. The effects reported in the graphs follow the coefficients found in the third model, including contextual peer effects and heterogeneous effects. The horizontal axis shows class composition as the proportion of indigenous students in class, divided in bins of 0.02. The vertical axis shows the change in score for every effect¹¹. Ethnic contextual peer effect curves are linear, since its effects depend linearly of the proportion of indigenous students. Since the aim of the graphs is to show the optimal level of class composition that would help indigenous students, the heterogeneous effects (the interaction term) was added as part of the ethnic contextual peer effect. Its addition does not change the sign of the effect, but does reduce its impact. Endogenous peer effects were calculated using the average of the class mean score for each bin of class composition. These effects are linear in terms of mean scores, but not in terms of class composition. The tendency of the effect is clearly positive, which reflects both the negative correlation between proportion of indigenous students and mean scores, and the negative effect of endogenous peer effects on individual scores. The red curve represents the total peer effects. This variable was constructed as a sum of the other two curves and all other contextual peer effects included

¹¹ The effect is measured as a change with respect to the mean of the variable. For example, the endogenous peer effects are measured as a change with respect to the average score, while the contextual peer effects of having indigenous in the class is measured as the change with respect to the average proportion of indigenous students in class.

in the analysis. The other contextual peer effects were found using the average mean characteristics of all students for each bin of class composition in terms of ethnicity.

Therefore, the total peer effects also incorporate the negative relation between proportion of rich students (as measured by high socioeconomic status) and the proportion of parents that have finished secondary education, and the proportion of indigenous students in class. The average number of male students does not have any strong correlation with the proportion of indigenous students in class.

It is possible to give an approximate optimal level of ethnic diversity in class by looking at where the total peer effects (red line) reach their maximum in the graph. For maths, the optimal proportion of indigenous students is around 60%. The optimal class composition for maximizing reading and history is around 88% and 66%, respectively. We should treat this only as the optimal level of ethnic diversity to maximize indigenous students' achievement, since the heterogeneous peer effects are incorporated¹².

Figure I



Note: Graphs show the effects of the third model.

¹² The optimal proportion of indigenous students in ethnic diverse classes for non-indigenous students will be the same without including the heterogeneous effects. Since these effects are small, we know the optimal level would be close, but below the one for indigenous students.

Two big assumption were made for this analysis. First, that the linear-in-means model is enough to evaluate each student's resulting score. This might not be the case, since many studies find non-linear effects and heterogeneities when evaluating peer effects (Lavy et al., 2012, Carrell et al., 2013). This means that, given the linear form of the model, it is susceptible to bias on the tails in the presence of non-linearities. The second assumption is that average observable characteristics of the class are fixed for each proportion of indigenous student in class. In other words, the average characteristics of the class (i.e. proportion of rich students in class) are directly associated to the proportion of indigenous students in class, and a change in proportion would imply a change in characteristics exactly as in the current data¹³. Therefore, we need to be careful in interpreting these results.

Given the limitations of the linear model, it is safer to treat the results as a linear approximation of a non-linear model, and constraint the analysis to values in the neighbourhood of the mean. The average proportion of indigenous students per class is 22%. It is clear that larger positive effects of ethnic diversity are found to the right of this mean. Therefore, even without finding the optimal level of ethnic diversity, it is clear that the mean is at a suboptimal point, and that an increase in the proportion of indigenous students per class would be desirable to improve their performance.

7. Conclusion, limitations, and policy implications

This study analyses the social interactions occurring in school classes shared by more than one ethnic group. The analysis was based on a linear-in-means model with fixed effects, first developed by Lee (2007). To the best of my knowledge, this study is the first to use Lee's model in the context of a developing country. It is also the first study that links the peer effects and ethnic stereotypes literatures in the context of the indigenous population of Peru. Additionally, it contributes to the model by adding interaction terms that allow for heterogeneities in the contextual peer effects.

The study successfully disentangles endogenous and contextual peer effects. The results shown prove that more than one channel exist, but also suggest that the channels are multiple and depend on other important characteristics like socioeconomic status. The reason behind

¹³ Adding indigenous students to a class will change all of the other observed average characteristics of the class.

this dependence might be the strong correlation between ethnicity, ability and these other characteristics.

One of the main limitations of this study is its inability to specify the causal mechanisms through which class composition affects individual results. However, the results do suggest possible mechanisms, like ability spill-overs, stigmatization, and heterogeneous dynamics within classes. The latter represents another important limitation of the study: If there are unequal interactions between students, the assumptions required for treating classes as reference groups would be violated which in turn would lead to the misinterpretation of the results. This is very likely to happen (and particularly worrying) in the presence of racial segregation. A way to solve this problem is to avoid choosing groups a priori, which requires the use of data that contains information about networks. On the other hand, it is possible that teachers are dividing their attention unequally across students, which would represent another case of heterogeneous interactions within class. If this is the case, incentives to focus the attention on disadvantaged students would be a good policy to reduce the gap in educational attainment. Another limitation of the model is that it does not account for heterogeneities in the endogenous peer effects. Since we are discussing racial stereotyping, it is possible that indigenous students will be more influenced by the results of their indigenous peers than by the results of others. Introducing heterogeneities to the endogenous peer effects remains an avenue for future research.

Part of the study was about the possibility of finding an optimal level of class composition that would improve indigenous students' scores. The study shows that class composition is a multi-faceted problem, and that differentiating between all peer effects is important to be able to reach the right class-composition policy. While it may not be possible to find an exact optimal class composition level, this study has shown that the level of ethnic diversity in the average Peruvian school is suboptimal if we want to improve indigenous students' attainment. It should be clear then that these students would benefit from a smaller proportion of non-indigenous peers.

8. References

- Agirdag O. et al. (2011). Ethnic school composition and peer victimization: A focus on the interethnic school climate. *International Journal of Intercultural Relations*, 35, 465–473.
- Ammermueller A, Pischke J-S. (2009). Peer effects in European primary schools: evidence from the Progress in International Reading Literacy Study. *Journal of Labor Economics*, 27(3): 315–348.
- Angrist, J. D., and Lang, K. (2002). Does school integration generate peer effects? evidence from boston's metco program. *The American Economic Review* 94, 5, 1613-1634.
- Aronson, J., Lustina, M.J., Good, C., Keough, K., Steele, C.M., and Brown, J. (1999). When white men can't do math: necessary and sufficient factors in stereotype threat". *Journal of Experimental Social Psychology*, 35(1), 29–46.
- Boucher, V., Bramoullé, Y., Djebbari, H., & Fortin, B. (2014). Do peers affect student achievement? Evidence from Canada using group size variation. *Journal of applied econometrics*, 29(1), 91-109.
- Bramoullé, Y., Djebbari, H., & Fortin, B. (2009). Identification of peer effects through social networks. *Journal of econometrics*, 150(1), 41-55.
- Buhs, E. S., Ladd, G. W., & Herald, S. L. (2006). Peer exclusion and victimization: Processes that mediate the relation between peer group rejection and children's classroom engagement and achievement? *Journal of Educational Psychology*, 98, 1–13.
- Carrell, S., Fullerton, R. & West, J. (2009). Does your cohort matter? Measuring peer effects in college achievement. *Journal of Labour Economics*, 27, 3.
- Carrell, S. E., Sacerdote, B. I. and West, J. E. (2013). From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation. *Econometrica*, 81: 855-882
- Cooley, J., (2009), Desegregation and the achievement gap: Do diverse peers help? *WCER Working Paper* No. 2008–7.

- Cueto, S., León, J. & Miranda, A. (2016). Classroom composition and its association with students' achievement and socioemotional characteristics in Peru. *Assessment in Education: Principles, Policy & Practice*, Vol. 23, No. 1, 126–148.
- Davezies L, d'Haultfoeuille X, Fougère D. (2009). Identification of peer effects using group size variation. *The Econometrics Journal* 12: 397–413.
- Felouzis, G. (2005). Ethnic segregation and its effects in middle school in France. *Revue Francaise De Sociologie*, 46, 3–35.
- Foster, Gigi (2006). It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism. *Journal of Public Economics*, Volume 90, Issues 8–9, 1455-1475.
- Glewwe, P., Krutikova, S., & Rolleston, C. (2014), Do schools reinforce or reduce learning gaps between advantaged and disadvantaged students? Evidence from Vietnam and Peru. *Working Paper 133*, Oxford: Young Lives.
- Gray-Little, B., & Hafdahl, A. R. (2000). Factors influencing racial comparisons of self-esteem: A quantitative review. *Psychological Bulletin*, 126, 26–54.
- Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science*, 11, 365–71.
- Kang, Changhui (2006). Classroom peer effects and academic achievement: Quasi-randomization evidence from South Korea. *Journal of Urban Economics* 61, 458–495
- Kudo, I., (2004). La educación indígena en el Perú. Cuando la oportunidad Habla Una Sola Lengua. In Donald R. Winkler, & Santiago Cueto (Eds.), *Etnicidad, raza, género y educación en América Latina*, 93–132. Santiago: PREAL.
- Lavy, V., Silva, O., & Weinhardt, F. (2012). The Good, the Bad, and the Average: Evidence on Ability Peer Effects in Schools. *Journal of Labor Economics*, 30(2), 367-414.
- Lee, L. F. (2007). Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics*, 140(2), 333-374.

Manski, C. F., (1993), Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*, vol. 60(3), 531-542.

Marx, D.M., & Roman, J.S. (2002). Female role models: protecting women's math test performance. *Personality and Social Psychology Bulletin*, 28(9), 1183-93.

Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics, Oxford University Press*, vol. 116(2), 681-704.

Steele, C.M., (1997). A threat in the air: how stereotypes shape the intellectual identities and performance of women and African Americans. *American Psychologist*, 52(6), 613-29.

Steele, C.M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797-811.

Valdivia, N. (2003). Etnicidad, pobreza y exclusión social: la situación de los inmigrantes indígenas en las ciudades de Cusco y Lima. In J. Uquillas, T. Carrasco y M. Rees (eds.), *Exclusión social y estrategias de vida de los indígenas urbanos en Perú, México y Ecuador* (pp. 29-129). Quito: Banco Mundial/Fideicomiso Noruego.

Valdivia, N., Benavides, M. & Torero, M. (2007). Exclusión, identidad étnica y políticas de inclusión social en el Perú: el caso de la población indígena y la población afrodescendiente. In *Investigación, políticas y desarrollo en el Perú* (pp. 603-655). Lima: Grade.

9. Appendices

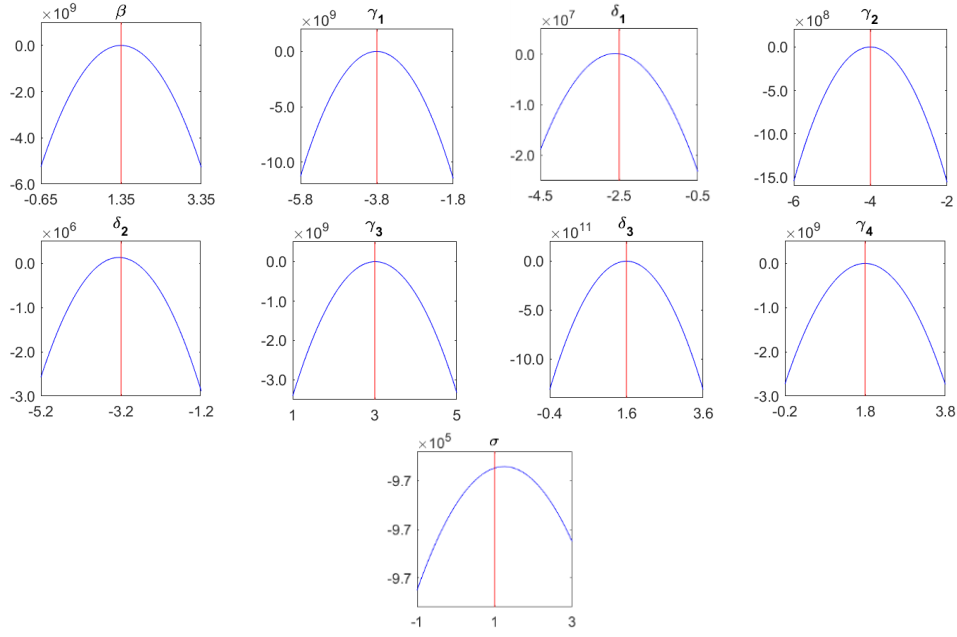
9.1. Simulating the CML estimator

The pseudo conditional maximum likelihood (CML) used for the estimations was coded using the computing software Matlab. The program maximizes the specified log-likelihood function for any given data. To test the performance of the estimator, data was simulated following equation (2), for randomly generated independent variables. Simulations were conducted for models with three independent variables. The main model added a fourth variable created as the interactions between one variable and its group mean. The simulation was replicated for different random coefficients. From now on, we will only refer to the full model of three different independent variables.

First, 500 classes were created. The size of each class was randomly generated following a normal distribution with a mean of 25 and a standard deviation of 7, in an attempt to replicate the real distribution of class sizes in the original data. Given the class sizes, the total number of observations in the simulated data is around 7000 students. The three variables (x_1 , x_2 and x_3) were generated following a uniform distribution, taking values of 0 and 1, and then randomly assigned to a class. A fourth variable was created as the interaction between x_3 and its group mean (remember: groups do not include individual i). Then, the differences between each variable and its class mean were created for each individual. Finally, the dependent variable y_2 was generated following equation (2), with a normal distributed error term with a mean of 0 and a standard deviation of 1. From the whole data, 2% of the variables were omitted to test how the model corrects for missing data.

The model was tested for different parameters, each of them randomly generated. Figure A1 shows the results for a CML estimator with three different independent variables, where β indicates the endogenous effect, γ_1 , γ_2 , and γ_3 are the individual effects for each of the three independent variables, and δ_1 , δ_2 , and δ_3 the contextual effects, for the same variables. γ_4 is the effect of the interaction term, which is considered an individual effect. As we can see in the figure, the likelihood function is well behaved and achieves its maximum close to the real parameters.

Figure A1: Shape of simulated Likelihood function



Note: Red line indicates the true parameters for the simulated data. This graphs show the effects for class size distributed as $N\{25,7\}$.

9.2 Ethnic segregation and determinants of harassment

One of the main argument of this paper is that one of the channels through which class composition affects indigenous students is through racial stereotypes and social segregation. Given that the data does not contain any way of measuring stigmatization, harassment was used as a proxy: It is clear that harassment and stigmatization are not equal, but we could expect a correlation between the two. Harassment is measured as a dummy variable, taking the value of 1 if the student reports had being a victim of either verbal or physical harassment by teachers or other students many times during the last year. Table IV presents 4 different regressions showing the relationship between harassment, ethnicity, socioeconomic status, gender, and class composition. Additionally, an interaction between class composition and the student's ethnic background designed to account for heterogeneous effects of class composition was added. The relationships are displayed using a simple OLS model.

Column 1, shows how having an ethnic background has a significant positive effect on the probability of being victim of harassment. In the second column, controls are added. Results show that being richer does significantly reduce the probability of harassment. However, ethnicity has a stronger correlation to harassment. The effect of having more indigenous students in class is positive and significant, but for indigenous students this effect almost counters the negative effect: Being in a class with students that share the same ethnic

background may be a good way to avoid harassment. Column 4 shows the same results, controlling for fixed effects at the school level. The effects of ethnicity and class composition become larger.

Table IV: Determinants of harassment

	1	2	3
Indigenous	0.082*** (0.003)	0.081*** (0.005)	0.100*** (0.005)
Socioeconomic Status (High)		-0.013*** (0.004)	-0.013*** (0.004)
Socioeconomic Status (Medium)		-0.008*** (0.003)	0.002 (0.003)
Male		-0.007*** (0.002)	-0.023*** (0.002)
Class composition (indigenous)		0.069*** (0.008)	0.148*** (0.020)
Indigenous*Class composition (indigenous)		-0.061*** (0.011)	-0.115*** (0.013)
Constant	0.239*** (0.001)	0.239*** (0.003)	0.223*** (0.005)
School Fixed Effects	No	No	Yes
Number of observations	141,578	141,578	141,578
R ²	0.006	0.007	0.086

Note: .01 - ***; .05 - **; .1 - *;