IDB
Inter-American
Development Bank

The effects of mass migration on the academic performance of native students. Evidence from Chile.

Migration Unit

Dante Contreras
Sebastian Gallardo

DISCUSSION
PAPER Nº
IDB-DP-816

June 2020

The effects of mass migration on the academic performance of native students. Evidence from Chile.

Dante Contreras
Sebastian Gallardo

June 2020

IDB
Inter-American
Development Bank

The effects of mass migration on the academic performance of native students. Evidence from Chile.*

Dante Contreras† Sebastián Gallardo‡

June 8, 2020

Department of Economics, Universidad de Chile

Abstract

Using longitudinal data and a Difference-in-Difference approach, this paper examines how the mass inflow of Venezuelan and Haitian migrants to Chile has influenced the learning outcomes of native students. The evidence shows negative effects on standardized test scores, especially in male students (-0.058SD). Said effects are considerably greater in Reading (-0.084SD) when examining the migration of non-Spanish speakers (mainly Haitians). When the shock is due to Venezuelan students, negative effects in Mathematics are observed equally strongly in male and female students (-0.053SD). Two factors that may account for these effects are identified. First, a lack of human and financial resources to address these new educational demands. Second, a native flight from public to private schools, especially involving higher SES students, when the shock is due to non-Spanish speakers (mainly Haitians).

Keywords: Immigration, Causal Effects, Academic Performance, Peer Effect.
JEL: F22, H52,I20,I28,J15

*We wish to thank Juan Blyde, Anna María Mayda, Matias Busso, and Juan Vargas, along with all the IDB migration workshop participants.

†Universidad de Chile, Department of Economics. Email: dcontrer@econ.uchile.cl

‡Universidad de Chile, Department of Economics. Email: segallardo@fen.uchile.cl

1 Introduction

Over the last decades, migration has become one of the most relevant topics around the world. We have seen a heated debate on the human rights of immigrants, the economic consequences of immigration, and public policies related to immigrants and their relationships with natives. In this context, the Latin America and the Caribbean (LAC) plays a major role in the world's migratory flows, as nearly 40 million migrants came from this region in 2017 (United Nations, 2017). Even though the favorite destination of LAC emigrants is North America, intra-regional migration emerges as the second preference, with 1 out of 6 migrants moving to a LAC country (Idem).

In this context, Chile has received one of the largest migrant flows over the last years. Between 2011 and 2018, migrant population increased nearly five-fold relative to the total population, jumping from 1.4% to 6.6%[1]. Since 2015, migratory flows to Chile have been strongly determined by foreigners coming from Venezuela and Haiti. This migration was motivated by political and economic difficulties in their countries of origin. Nowadays, these groups constitute the first and third largest foreign communities in the country. As a result, Chile provides an interesting opportunity for empirically examining the consequences of a rapid mass migration phenomenon.

The empirical literature that has examined the effects of migration has typically focused on the labor market outcomes of native workers[2], but there also exists a growing literature focused on educational outcomes, specifically on the academic performance of native-born students. This literature has reported mixed results. For instance, Diette and Uwaifo Oyelere (2014) explore the peer effects of Limited English students on the Reading and Math grades of native-born school students, differentially by gender and race. Using 1998-2006 administrative data of 4th to 8th grade students from North Carolina and a school-by-year fixed effect approach, the authors find zero effect on native female students, both white and black. However, they also report a modest negative effect on males, particularly on black males' scores in both subjects and on white males' math scores. In a previous paper, using the same data and similar specifications, these authors had provided evidence of heterogeneous effects on native-born students (as reflected by their Math and Reading test scores). Positive effects are found in the bottom and middle parts of the performance distribution, while small negative effects are found at the top (Diette and Uwaifo Oyelere, 2012).

[1] Data for 2011 were obtained from the CASEN survey of that year, while data for 2018 were provided by the National Statistics Institute.

[2] Okkerse (2008), Longhi et al. (2005), and Dustmann et al. (2016) provide good literature summaries, while at the same time outlining the main challenges and empirical differences involved in identifying the effects of migration. Contreras and Gallardo (2020) provide empirical evidence of the impact of recent migratory flows in Chile on native wages.

In another study, Conger (2015), using administrative records from the Florida Department of Education, evaluates the effects of different shares of immigrant English language learners (ELL) and non-English language learners (non-ELL)[3] on four cohorts of 9th graders attending Florida public high schools between 2000 and 2004. Using fixed effects by school and cohort, results indicate that immigrant peers have no effect on their classmates' academic achievement.[4]

In contrast, Jensen and Rasmussen (2011), using matched administrative records and PISA results of 9th graders from Denmark in 2000, found a negative effect of immigrant concentration in schools on both Reading and Math. However, the authors acknowledge that these results are only robust with a IV approach for Math.

Figlio and Özek (2019), using administrative schooling and birth records, examine how Haitian refugees to Florida after the 2010 earthquake in Haiti influenced the educational outcomes of 3th to 10th grade incumbent students in public schools. Adopting different specifications to address the non-random sorting of refugees within the Florida school system, they found very modest positive effects on the academic performance of native students (particularly in Reading) the year of the earthquake and two years later.

This study aims to identify the effect of recent mass migratory influxes on the academic performance of native-born Chilean students. To do so, we take advantage of the quasi-experimental situation generated by the migratory shocks from Venezuela and Haiti between 2016 and 2018. Using census and longitudinal data from the SIMCE test[5] provided by the Ministry of Education's Education Quality Agency and the General Student Information System, we applied a Difference-in-Difference strategy on the change in SIMCE scores between 4th and 6th grade. Native students whose schools received migrants from Venezuela, Haiti, or other non-Spanish-speaking countries in 2018, after having none in 2016, were the treatment group, while those who had no peers from these countries during this period were the control group. Both groups were balanced in terms of their observable characteristics using a nearest neighbor matching strategy with no replacement. In addition, we employ the same strategy to evaluate the effect separately, that is, when the shock affecting a school was due to Venezuelan students only or when it was exclusively produced by non-Spanish-Speaking students, largely driven by influxes of Haitian peers.[6] Results in-

[3] An ELL student is one who falls below a predetermined score on English proficiency tests.

[4] The sample used to present the main results in this paper comprises all students and not only native-born ones. In the online appendix, the author restricts the sample to native-born students, again finding a zero effect.

[5] SIMCE is the standardized test score which measures educational achievement in several subject areas. Section 3.1 provides a detailed explanation of the data.

[6] Since the flow of immigrants from Haiti is concentrated in specific regions of the country, we decided to examine them within a larger category but in accordance with the shock that they generate on native-born students (i.e. the shock generated by students who do not speak Spanish as their mother tongue). In any

dicate that the shock caused by the arrival of migrant students between 2016 and 2018 had a negative effect on the SIMCE Reading scores of male native-born students between 4th and 6th grade (-0.058SD). This effects appears to be greater when the shock is due to influxes of migrants from non-Spanish-speaking countries (-0.084SD). In mathematics, both male and female students exposed to the shock of Venezuelan migrants performed more poorly between 4th and 6th grade compared to the controls (-0.053SD). Results suggest that there are insufficient human and financial resources to address these new educational demands in the region. Indeed, it can be observed that the student/teacher ratio increases due to the shock, which is consistent with the idea that limited resources are the mechanism behind the results obtained. The Chilean educational systems allocates resources to schools according to student enrollment, but does not grant any additional resources to those that serve immigrant students.

In addition, when observing the shock due to non-Spanish-speaking students, we find evidence of native flight of higher SES students from public to private schools (fully private and State subsidized). This apparent cream-skimming effect may be an additional mechanism leading to the negative effects observed in the academic performance of native-born students.

This study contributes to the literature on the peer effect of migrant students in the classroom on various educational outcomes. First, to the best of our knowledge, this is one of the first studies to show evidence about immigrant peer effect in LAC. Second, by taking advantage of the quasi-experimental situation generated by the migratory influxes of recent years, we provide causal evidence of the short-term classroom effects of migrant students flows, considering that native-born students had no foreign classmates beforehand. Third, we provide a more in-depth examination of the mechanisms that may be influencing outcomes, such as greater restrictions on school resources due to the greater student/teacher ratio and the native flight of higher-SES native-born students. Lastly, the quality of the longitudinal and administrative census data in this study supports the accuracy of the results found.

After this introduction, section 2 provides statistics to contextualize the recent situation of migration within the Chilean educational system. Then, in section 3, we explain the data and the strategy to be used to evaluate the effect of recent mass migratory inflows on the academic performance of native Chilean students, along with the possible mechanisms behind them. The fourth section shows our main findings, while the fifth and last section contains our conclusions.

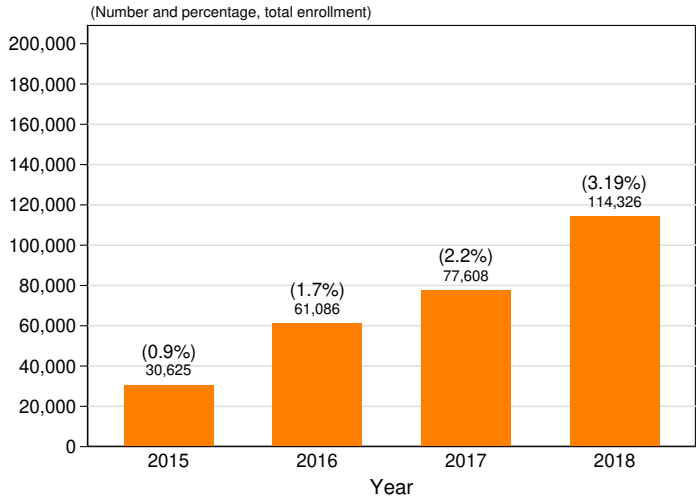
case, the main results hold if we only consider the shock caused by Haitian students, but statistical power is reduced due to the smaller number of observations.

2 Current context of migration in the Chilean educational system

In Latin America and the Caribbean, Chile is one of the countries that has received the most immigrants over the last years. Figure 1 reveals a major increase in the number of children born in another country who have enrolled in the Chilean school system. According to information issued by the Research Center of the Ministry of Education (Mineduc) the number of immigrant children has increased 3.5 times between 2015 and 2018, with students from other countries reaching 114,000 .

This trajectory is also characterized by the distribution of immigrants by country of origin. Figure 2 shows the progression in the number of migrant students in the Chilean school system for the six main communities from 2015 to 2018. The period is marked by the rapid increase in enrollees from Venezuela and Haiti since 2016, with students from these countries becoming the first and second colony in Chilean schools in 2018. In contrast, the number of Peruvian, Colombian, and Bolivian students is growing but at a relatively slowly pace. In fact, the Peruvian student community, once the largest colony in Chile, was surpassed in the last year under study by Venezuela and Haiti as previously noted.

Figure 1: Evolution of the number and percentage of migrant students in Chile, 2015-2018.



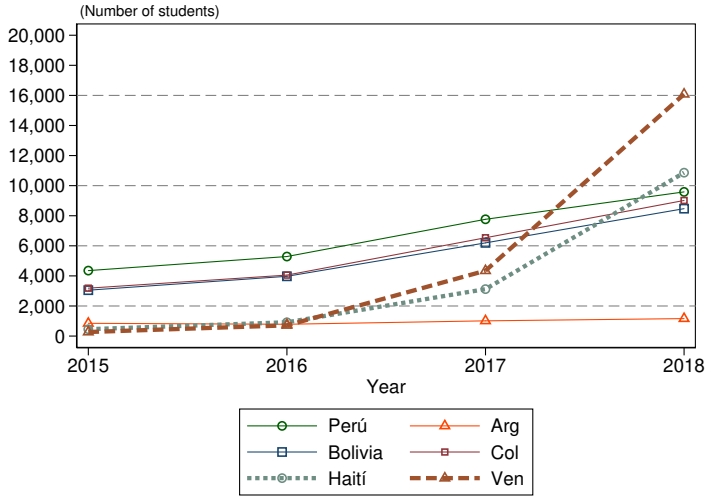
Source: Sistema General de Estudiantes (SIGE), 2015-2018.

Given Chile's particular geography, the regional distribution of migrants is another relevant consideration. Table 1 shows the fraction of migrant students within the entire Chilean educational system along with the number and distribution of migrant students across

all regions of Chile. The table reveals that the fraction of migrants between regions is significantly heterogeneous. The Tarapacá Region stands out in this regard, as migrants constitute 10.4% of its enrollment. This region is followed by the Antofagasta region (10%), the Arica and Parinacota Region (7%), and the Metropolitan Region (5.1%). However, it should be noted that over 60% of migrant students reside in the Metropolitan Region.

Table 2 shows the distribution of foreign students across the levels of the educational system.[7] It can be observed that most of these students attend primary school; within this level, they are quite evenly distributed among the first 8 grades. In contrast, migrant students are much less prevalent in secondary education, with their number becoming smaller in higher grades.

Figure 2: Evolution of the number of immigrant students in Chile by country of origin, 2015-2018.



Source: Sistema General de Estudiantes (SIGE), 2015-2018.

Lastly, the first column of Table 3 shows the distribution of native-born students in the three types of school that exist in Chile, with most of them attending State-subsidized private schools. In contrast, the second column shows that most migrant students attend public schools. More specifically, the third column shows that more than 60% of Venezuelan students are concentrated in public schools, while the fourth column indicates that over two thirds of Haitian students and those from non-Spanish-speaking countries

[7] These data only refer to non-adult students. The categorization of adult students, who account for less than 4% of the sample, is more complicated.

attend public institutions. This is due to the fact that public institutions, unlike State-subsidized private ones, do not conduct selection processes and provide education free of charge; therefore, they can be more easily accessed by migrant students, who tend to come from socioeconomically disadvantaged backgrounds.

In summary, recent statistical reports indicate that the number of students born in other countries and who enroll in the Chilean school system has increased significantly. This enrollment exhibits a high level of heterogeneity across regions and school administrative dependence. However, this enrollment pattern is rather homogeneous across grades, especially in the primary system. Thus, it is important to study how this experience has influenced student performance in the educational system.

Table 1: Regional statistics of immigrant students (2018).

Region	**Total**	**Fraction**	**Distribution**
Region I: Tarapacá	8,289	0.104	7.25%
Region II: Antofagasta	13,154	0.100	11.51%
Region III: Atacama	1,763	0.027	1.54%
Region IV: Coquimbo	2,561	0.015	2.24%
Region V: Valparaíso	4,976	0.014	4.35%
Region VI: O'Higgins	2,625	0.014	2.30%
Region VII: Maule	2,008	0.009	1.76%
Region VIII: Biobío	2,290	0.005	2.00%
Region IX: Araucanía	852	0.004	0.75%
Region X: Los Lagos	1,063	0.006	0.93%
Region XI: Aysén	223	0.009	0.20%
Region XII: Magallanes	629	0.020	0.55%
Region XIII: Metropolitana	69,919	0.051	61.16%
Region XIV: Los Ríos	305	0.004	0.27%
Region XV: Arica y Parinacota	3,669	0.070	3.21%
Total	114,326	0.031	100%

Source of migrant student data: Sistema General de Estudiantes (SIGE), 2018.

Table 2: Statistics of immigrant students by grade.

Primary				Secondary			
Grade	Total	Fraction	Distribution	Grade	Total	Fraction	Distribution
1st	8,869	0.035	13.56%	9th	7,572	0.031	31.88%
2nd	8,572	0.033	13.11%				
3rd	8,430	0.033	12.89%	10th	6,741	0.029	28.38%
4th	8,080	0.032	12.35%				
5th	7,829	0.032	11.97%	11th	5,542	0.026	23.33%
6th	7,793	0.032	11.92%				
7th	7,875	0.033	12.04%	12th	3,896	0.019	16.40%
8th	7,955	0.034	12.16%				
Total	65,403	0.033	100.00	Total	23,751	0.026	100.00

Source: Sistema General de Estudiantes (SIGE), 2018. The sample is restricted to non-adult students.

Table 3: Distribution of natives and migrants by school type (2018).

	Natives	Migrants	Venezuelans	Haitians and non-Spanish speakers
Public	36.1%	60.3%	61.1%	67.5%
Subsidized private	54.1%	32.6%	35.7%	28.7%
Private	9.8%	7.0%	3.2%	3.9%

Source of migrant student data: Sistema General de Estudiantes (SIGE), 2018.

3 Methodology

3.1 Data

To study the effects of mass immigration on the academic performance of native-born students, we will use SIMCE panel data to follow the same students in 2nd, 4th, and 6th grade in 2014, 2016, and 2018 respectively. This database derives from a census-type measurement performed annually by the Education Quality Agency at several levels of the primary and secondary system. It comprises individual sociodemographic variables such as parents' educational level and the income level of students' families. Also, it contains school variables like school type (public, subsidized, or private), whether the school is rural, the region where the school is located, and the socioeconomic background of the institution.[8] Finally,

[8]This variable has 5 categories: low, medium low, medium, medium high, and high. It is constructed considering parents' years of schooling and income level as reported in the SIMCE parents' survey, at the school and grade level.

this database contains each student's scores in the Mathematics and Reading SIMCE test.

The General Student Information System (SIGE) database is a set of census data and administrative records generated by MINEDUC's Research Center. This database reveals whether a student was born in Chile or abroad, making it possible to impute this information with SIMCE data.[9] This database contains the information of all students in the Chilean school system, thus enabling us to calculate the number of migrants at the grade and school level for each year from 2015 to date.

The final sample comprises all students who remain in the same school[10] in 2014, 2016, and 2018, from 2nd to 6th grade. This makes it possible to observe the academic performance variables of the native students who have stayed in the same school during the first years of primary education and determine how their performance is affected by sharing their classroom with students arriving from another country, specifically Venezuela and Haiti or other non-Spanish-speaking countries. We selected the cohort of students who were in 2nd grade in 2014 because they have taken the SIMCE test every two years, which makes it easier to identify effects between these two periods.[11]

3.2 Identification Strategy

To identify the effect of the recent flows of foreign students on the academic performance of native students, we used a Differences-in-Differences (DD) strategy, represented by the following equation:

$$SIMCE_{ist} = \alpha T_{is} + \delta POST_t + \beta T_{is} * POST_t + X_{ist}\gamma + \varepsilon_{ist} \quad (1)$$

To examine the effects of the mass migrant inflow of recent years, we conduct an analysis using three samples. First, we examine the shock caused by the whole migratory inflow. Second, we examine the effects of Venezuelan migration only. Lastly, we examine the effects of the shock associated with non-Spanish-speaking immigration.

In equation (1), the dependent variable is the standardized SIMCE score on the Reading or Mathematics test of student i attending school s during period t. Variable T indicates the treatment: 1 if student i had no classmates from Venezuela, Haiti, or another non-Spanish-speaking country in 2016, when he/she was in 4th grade in school s, but had at least one classmate from one of these countries in 2018 when he/she was in 6th grade in

[9] The same data source provides the GPA score in 2nd and 4th grade for the placebo exercise described in the following subsection.

[10] This restriction is eliminated when studying the native flight effect, described in section 4.3.

[11] Other cohorts, such as 6th grade students in 2014, did not take the test until 10th grade (2018), so performing analyses with this cohort, with multiple untested grades and years in between, can lead to spurious results.

the same school s. This variable assigns value 0 to student i if he/she has no classmates from any of these countries in both periods. Variable $POST$ is 1 in the year after the migrant shock, that is, in 2018, when the native student is in 6th grade, and 0 for the year 2016, when the native student is in 4th grade. Vector X contains controls such as student sex, father's and mother's educational level, household income level, the region where the school is located, school type (public, private, subsidized), whether the school is in a rural area, and the socioeconomic group to which the school belongs. Finally, ε is the error term clustered at the student level.

To evaluate the effect of the shock caused by the increase in Venezuelan students, the definition of treatment T considers that native student i was exposed to the shock if, in 2016 (4th grade), he/she had no classmates from Venezuela in school s, but had at least one in 2018, when he/she was in 6th grade in the same school; the value is 0 if he/she had no Venezuelan classmates during this period.

Likewise, to evaluate the effect of the shock due to the inflow of Haitian students, variable T takes the value 1 if native student i who was in 4th grade in 2016 had no classmates from a non-Spanish-speaking country in school s, but had at least one in 2018, when he/she was in 6th grade in the same school; the value is 0 if he/she had no classmates from these countries throughout this period. Since Haitian students are not present in all regions or school types in 6th grade, the category is extended to non-Spanish-speaking students, where Haitians represent nearly 90% of the total. This process yields observations in all regions and school types, which preserves the representation of these two variables in our estimates and prevents the loss of statistical power.[12]

Thus, the β coefficient is the Differences-in-Differences estimator, which captures the causal effect of the migratory shock of Venezuelan and Haitian students[13] on the academic performance of native-born students as measured through standardized SIMCE scores, compared to native students who were not exposed to this shock in their classrooms between 4th and 6th grade.

Given that the treatment group has a much smaller number of observations than the control group, differences in observable covariates are relevant. Therefore, we conducted a process based on propensity score matching using nearest neighbor search, with no replacement.[14]

[12] Figure 3 in the Appendix shows the number of Venezuelan, Haitian, and non-Spanish-speaking students in 6th grade. Please note that there are no 6th grade Haitian students in several regions, but after extending the category to non-Spanish-speaking students, all regions except for Aysén have at least one observation. The same is true of private schools, which have no Haitian students. In any case, we control for region and school type to avoid any bias due to the construction of this category.

[13] Venezuelan (1,101) and non-Spanish-speaking students (539) in 6th grade represent slightly over 20% of the total number of migrant students in this grade in 2018.

[14] Nearest neighbor matching with no replacement achieved the best balance between the control covariates

In addition, we applied the estimation described in (1) to the SIMCE Reading score in 2nd and 4th grade between 2014 and 2016 as a placebo exercise intended to show the validity of the DD results. To do so, we followed exactly the same native students placed in the treatment or control groups, according to the previous definitions; however, we applied the strategy considering the *POST* variable to be 0 if the student took the SIMCE test in 2014, when he/she was in 2nd grade, and 1 if he/she took it in 2016, when he/she was in 4th grade. Since the 2nd grade SIMCE test does not cover Mathematics, we used the GPA score in 2nd and 4th grade as a proxy for the Mathematics score.

Finally, in order to evaluate possible mechanisms through which results influence academic performance, we will apply the DD of equation (1) to the student/teacher ratio and a Triple-Differences (DDD) strategy to the likelihood of a native student i attending a subsidized or private school, with the third interaction including a dummy that indicates whether he/she belongs to the highest decile of household income. These estimates are explained in more detail in the Results section.

3.3 Descriptive Statistics

Tables 9, 10, and 11 (see Appendix) show the descriptive statistics of the treatment and control groups considering the Venezuelan migrant shock and the non-Spanish-speaking migrant shock, along with the separate shocks for the baseline period (2014).

The tables reveal that, despite the use of propensity score matching, some marginal differences remain between the groups. According to Table 9, which shows the means of the observable covariates of both groups (treatment and control, with 22 thousand observations each) related to the Venezuelan or Haitian shock[15], differences in representation exist only in the highest household income category (above CLP$2MM), in the O'Higgins region, in private schools, and in the two extreme school socioeconomic background categories. In contrast, Table 10 shows the means of the observed variables in the groups of students who were not exposed to a Venezuelan migrant shock in 4th or 6th grade, compared to those who were exposed (with both groups comprising over 19,000 each). Girls are more extensively represented in the treatment group and only one income level category (between CLP$200,000 and CLP$300,000) differs between the two groups. Similarly, the distribution of students among these groups only differs in two regions (O'Higgins and Bío-Bío). Lastly, the treatment and control groups differ only in the first three school socioeconomic background categories (the lowest ones). Even though none of these differences are of a

in the treatment and control groups, surpassing other strategies that involved matching with more neighbors and replacement.

[15] Henceforth, this shock will be referred to as "full migrants shock"

relevant magnitude, it is worth controlling for them in estimates.

Meanwhile, Table 11 shows that the students subjected to the inflow of students from non-Spanish-speaking countries strongly resemble those not affected by this shock (some 3,800 students in each group) in terms of most of their observable covariates. In fact, marginally significant differences are observed in the distribution of only one region of the country (O'Higgins).

Finally, Figures 4, 5, and 6 show the histograms of the common support between the treatment and control groups after the application of propensity score matching (panel a) and the comparison between the propensity scores estimated between both groups, in ascending order by propensity score (panel b), for each of the shocks analyzed.

4 Results

4.1 The effect of immigration on academic performance

The top panel of Table 4 shows the results of the DD estimation of equation (1). The first column indicates that native students subjected to a shock caused by the presence of migrant students in their classroom between 2016 and 2018 performed more poorly on the SIMCE Reading test (0.045SD) than native students who received no foreign students in their school and grade. Columns 2 and 3 disaggregate the results of the first column, but restricted to boys and girls respectively. The effect on boys is a 0.066SD reduction in their SIMCE Reading score; however, in girls this negative effect is nearly 3 times milder, reaching -0.028SD. The last three columns of the same table show the effect of the shock caused by migrants from Venezuela and Haiti or non-Spanish-speaking countries on SIMCE Mathematics scores. Here, the effect is again greater in boys, with a score reduction of 0.051SD, compared to a 0.041SD reduction in girls.

Finally, panel B of Table 4 shows the results of the placebo exercise used to validate the results of the DD strategy presented in panel A. To do this, we imputed the treatment condition obtained when evaluating the 2016-2018 period to the same students, but for the 2014-2016 period. Thus, in the placebo exercise, we analyzed exactly the same students from the treatment and control groups, but now the $POST$ variable is 1 if the SIMCE test was taken in 2016, when the student was in 4th grade, and 0 if it was taken in 2014, when the student was in 4th grade. For the placebo of Mathematics scores, we evaluated the change in GPA between 2nd and 4th grade, since 2nd grade SIMCE only has a Reading section. We know that there was no migrant influx shock in this period; therefore, the coefficients of the DD estimation should be null. This, was indeed the case, thus validating the DD strategy applied.

Panel A of Table 5 shows the results of the DD estimation of equation (1), which identifies the shock due to Venezuelan students. The first column indicates that the native students who were affected by the inflow of migrant students from Venezuela in their classroom between 2016 and 2018 performed more poorly on the SIMCE Reading test (approximately -0.04SD) than the native students who received no students from that country in their school and grade. Columns 2 and 3 disaggregate the results of the first column for boys and girls respectively. The effect found in the full sample is driven by boys only, with the DD coefficient being -0.058SD for the SIMCE Reading score. Although the effect on girls is negative as well, it is of a much smaller magnitude and non statistically different from zero.

Table 4: Effects of all student migration shock on the SIMCE standardized scores of native students between 4th and 6th grade (2nd and 4th - placebo test).

	Reading			Math		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Full	Boys	Girls	Full	Boys	Girls
Panel A: *4th to 6th*						
Difference-in-Difference	-0.045***	-0.066***	-0.028*	-0.045***	-0.051***	-0.041***
	(0.010)	(0.015)	(0.014)	(0.010)	(0.015)	(0.014)
Observations	42,020	20,509	21,511	42,221	20,581	21,640
Number of students	23,563	11,568	12,012	23,611	11,581	12,046
Overall R-squared	0.102	0.0980	0.0992	0.142	0.148	0.135
Panel B: *2th to 4th (placebo test)*						
Difference-in-Difference	-0.012	-0.008	-0.016	-0.003	-0.007	0.000
	(0.011)	(0.017)	(0.016)	(0.005)	(0.007)	(0.006)
Observations	42,545	20,750	21,795	43,576	21,364	22,212
Number of students	23,640	11,607	12,065	23,957	11,812	12,177
Overall R-squared	0.107	0.108	0.101	0.114	0.114	0.0972

Note: Clustered standard errors at student level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The sample is limited to native students who did not switch schools between 2014 (when they attended 2nd grade) and 2018 (when they attended 6th grade). The control variables of the estimations are gender, years of schooling of both parents, and the student's household income level. We also include a categorical variable for the region where the school is located, a categorical variable for school type (public, private, subsidized) and socioeconomic background, and a rural school dummy. In panel B, the placebo Reading test is the change in standardized SIMCE Reading score between 2nd and 4th grade (first 3 columns) and the placebo Mathematics test is the change in GPA between the same grades (columns 4 to 6).

Column 4 indicates that native students exposed to the shock of Venezuelan student inflow scored more poorly on the SIMCE Mathematics test between 4th and 6th grade (-0.053SD) compared to native students not exposed to the same shock. Lastly, the fifth and sixth columns do not reveal a large difference between the estimates for boys and girls; that is, both performed more poorly on the SIMCE Mathematics test (approximately -0.053SD).

Table 5: Effects of Venezuelan student migration shock on SIMCE standardized scores of native students between 4th and 6th grade (2nd and 4th - placebo test).

	Reading			Math		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Full	Boys	Girls	Full	Boys	Girls
Panel A: *4th to 6th*						
Difference-in-Difference	-0.037***	-0.058***	-0.019	-0.053***	-0.054***	-0.053***
	(0.011)	(0.017)	(0.016)	(0.011)	(0.016)	(0.015)
Observations	35,009	17,128	17,881	35,161	17,177	17,984
Number of students	19,611	9,641	9,983	19,647	9,649	10,010
Overall R-squared	0.101	0.0964	0.101	0.142	0.144	0.137
Panel B: *2th to 4th (placebo test)*						
Difference-in-Difference	-0.006	0.011	-0.022	-0.003	-0.001	-0.005
	(0.012)	(0.018)	(0.017)	(0.005)	(0.008)	(0.007)
Observations	35,477	17,353	18,124	36,239	17,792	18,447
Number of students	19,676	9,674	10,026	19,910	9,824	10,110
Overall R-squared	0.107	0.108	0.103	0.118	0.116	0.105

Note: Clustered standard errors at student level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Same details as Table 4

The bottom panel of Table 5 shows the results of the placebo exercise. For Reading, the results of the DD coefficient are quite minor compared to the upper panel and statistically non-significant. The results of the placebo on GPA (last three columns of panel B) also validate our findings for Mathematics in the upper panel for the full sample and for boys and girls separately.[16] Therefore, these results are robust to the placebo test.

The upper panel of Table 6 shows the results of the estimates based on equation (1) when the shock is due to the influx of migrants from non-Spanish-speaking countries into a native student's school between 2016 and 2018. Column 1 indicates that students exposed to this shock scored more poorly on the SIMCE Reading test (-0.055SD) between 4th and 6th grade. This effect is concentrated in the boys subsample (column 2), with an effect of -0.084SD, and a negative but statistically non-significant effect among girls.

The last 3 columns of Table 6 reveal that negative but statistically non-significant results were observed in all cases for the SIMCE Mathematics test.

Lastly, the bottom panel of Table 6 shows the results of the placebo exercise, which are similar to those presented in Table 5. The coefficients associated with change in the SIMCE

[16]However, Mathematics scores must be cautiously considered because the placebo exercise was not conducted with the variable that was originally examined (SIMCE Mathematics).

Reading test between 2nd and 4th grade (columns 1 to 3) are all statistically non-different from 0 and of a very small magnitude, which is similar to our analysis of GPA as a proxy for the Mathematics test (columns 4 to 6).

In brief, the arrival of migrant students between 2016 and 2018 had a negative effect on the SIMCE Reading scores of native-born boys between 4th and 6th grade. This effect is greater when it is due to inflows of non-Spanish-speaking migrants compared to the Venezuelan student shock. In Mathematics, both girls and boys exposed to the Venezuelan migrant shock performed more poorly between 4th and 6th grade compared to the control group.

Table 6: Effects of the inflow shock of Non-Spanish-speaking immigrant students on the standardized score of native students between 4th and 6th grade (2nd and 4th grade - placebo test).

	Reading			Math		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Full	Boys	Girls	Full	Boys	Girls
Panel A: *4th to 6th*						
Difference-in-Difference	-0.055***	-0.084***	-0.028	-0.020	-0.020	-0.021
	(0.018)	(0.028)	(0.025)	(0.018)	(0.026)	(0.024)
Observations	13,524	6,375	7,149	13,592	6,388	7,204
Number of students	7,689	3,656	4,039	7,708	3,660	4,054
Overall R-squared	0.111	0.110	0.103	0.151	0.161	0.143
Panel B: *2th to 4th (placebo test)*						
Difference-in-Difference	0.001	-0.001	0.003	-0.006	-0.011	-0.001
	(0.021)	(0.031)	(0.028)	(0.009)	(0.014)	(0.012)
Observations	13,768	6,491	7,277	14,172	6,716	7,456
Number of students	7,716	3,669	4,058	7,835	3,743	4,103
Overall R-squared	0.111	0.115	0.102	0.115	0.107	0.102

Note: Clustered standard errors at student level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Same details as Table 4.

4.2 Student to teacher ratio as a mechanism

In this section, we evaluate if limited school resources can operate as a mechanism behind the results presented in the previous section. To do this, we will use the student to teacher ratio as a proxy for the resources available to a school in each period. This strategy is based on the idea that, if this ratio increases due to inflows of migrant students, schools will have less resources for each student on average, which may lead to a performance reduction compared to schools that received no migrant students. This approach is similar to that used by Farre et al. (2018). Our dependent variable was the student to teacher ratio at the

school level in primary education. This indicator is constructed as the ratio of the total number of primary education students in each school to the number of teachers in this level. As shown in section 2 (Table 2), migrant students are not unequally distributed among primary school grades, which means that specific grades cannot be receiving a larger migrant inflow. Therefore, this measure is representative of all primary grades and –logically– of each grade studied in our estimates. We use each school as the unit of observation, while the values of the control variables are equal to the average value of the native students of each grade studied (2nd, 4th, and 6th grade, depending on the year observed).

The results of estimating[17] equation (1) using the student to teacher ratio are shown in Table 7. Panel A of Table 7 shows the effect of the joint Venezuelan and Haitian shock, indicating a significant increase in the student to teacher ratio. This parameter decreases when controls (0.67 to 0.53) are added in column 2. The results that summarize the placebo effect are shown in columns 3 and 4. Such effects are statistically equal to zero (column 4) or marginally significant (column 3).

Panel B of the table shows that, when the shock is due to students from Venezuela, there is a statistically significant increase in the student to teacher ratio. The estimated coefficient becomes quite robust after adding all the control variables (second column). The third and fourth columns show the results of the placebo exercise with and without controls. In line with the results presented in the previous section, the placebo exercise validates the DD strategy with the treatment and control groups, since no significant changes are observed in the student to teacher ratio in the period before the immigrant student shock.

Panel C of Table 7 shows the effect on the ratio when the shock is due to students from non-Spanish-speaking countries. Again, the effect is positive, significant, and robust to the inclusion of the control variables and the placebo test. Yet, it should be noted that the coefficients estimated are slightly higher than those shown in Panels A and B.

All in all, even though the schools in the treatment group employ their available resources more intensively (using the student to teacher ratio as a proxy) than the control group, this effect appears to be quite moderate; for instance, it does not result in a one-student increase in the ratio, suggesting that it is not the main mechanism behind the reduction in the academic performance of the treatment group.

School funding in Chile is another element that should be considered when interpreting these results. Schools receive funds according to the number of students enrolled in them; therefore, more students should lead to additional funds. However, this increase should

[17] Treatment and control groups were allocated using the same schools placed in each category after the nearest neighbor matching method applied in the previous section.

be compared with the rising costs of integrating new students. For instance, for a long time, students with special educational needs received a per-child subsidy similar to that of a regular student. Furthermore, there was no accurate information identifying these children's needs. Similarly, per-student vouchers were flat. A vulnerable child and another from a higher SES background received the same amount. A number of policies were implemented to correct these issues. Basically, resources were increased to cover the costs of working with children whose education entailed more complexity. Evidence shows that these additional resources led to improvements in these children's learning outcomes[18].

With respect to migrant children, there are no additional subsidies that reflect the potentially higher cost of their education. Certainly, tasks such as integrating them, bringing them up to their peers' academic level, arriving at a common terminology, and providing additional coaching in subject areas such as History or Language, among others, entail additional costs that should be weighed in order to attain educational integration.

[18] See Contreras et al. (2020) for an analysis on special educational needs and Nielson (2013) and Navarro-Palau (2017) for SEP.

Table 7: Effects of the immigrant student shock on student to teacher ratio.

VARIABLES	(1)	(2)	(3)	(4)
	4th to 6th		2nd to 4th (*placebo test*)	
Panel A: *All migrants shock*				
Difference-in-Difference	0.668***	0.534***	-0.267*	-0.231
	(0.148)	(0.149)	(0.153)	(0.156)
Observations	6,646	6,631	6,646	6,609
Number of schools	3,324	3,324	3,324	3,324
Overall R-squared	0.00797	0.261	0.00573	0.247
Panel B: *Venezuelan shock*				
Difference-in-Difference	0.781***	0.613***	-0.254	-0.212
	(0.168)	(0.170)	(0.168)	(0.172)
Observations	6,032	6,025	6,097	6,057
Number of schools	3,049	3,049	3,049	3,049
Overall R-squared	0.004	0.256	0.003	0.235
Panel C: *non-Spanish-speaking shock*				
Difference-in-Difference	0.811***	0.749***	-0.228	-0.199
	(0.226)	(0.228)	(0.257)	(0.261)
Observations	3,784	3,779	3,784	3,761
Number of schools	1,893	1,893	1,893	1,893
Overall R-squared	0.001	0.237	0.002	0.218
Controls	NO	YES	NO	YES

Note: Clustered standard errors at school level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The control variables of the estimations (column 2) are students' gender, years of schooling of both parents, and household income level (calculated as an average at the school-grade level in each period). We also include a categorical variable of the region where the school is located, a categorical variable for school type (public, private, subsidized) and socioeconomic background, and a and a rural school dummy.

4.3 Native flight and cream-skimming effect as mechanisms

Regarding this topic, authors have often analyzed the flight of local students from public to private schools due to the the increasing number of migrant students in public schools, a phenomenon known as *native flight effect*. When it is higher SES students who transfer to private schools, authors speak of a *cream-skimming effect*. As students with more economic resources at home tend to perform better academically, their enrollment in other schools may reduce the benefits of the peer effect that they can generate, especially among lower SES students.

To measure *native flight* and the *cream-skimming effect*, we developed the following Triple-Differences (DDD) strategy [19]:

$$PRIVATE_{it} = \beta_1 T_{is} * POST_t + \beta_2 T_{is} * POST_t * HighSES_{is} + X'_{ist}\phi + \nu_{ist} \quad (2)$$

The $PRIVATE_{it}$ variable is a dummy that takes the value 1 if student i is enrolled in a private or subsidized private school in t and 0 if the student attends a public school[20]. In the right-hand part of equation (2), coefficient β_1 captures the native flight effect (DD coefficient) and β_2 captures the cream-skimming effect (DDD coefficient). The variable T_{is} represents one of the three treatments that have been separately analyzed in the previous estimations. It should be noted that, in the DDD interaction, we add the dummy variable $HighSES_{is}$, which indicates whether the student belongs to the highest household income decile in the sample.[21] The vector X'_{ist} contains the same control variables as equation (1), but along with all the other interactions derived from the DDD estimation. The term ν_{ist} represents cluster errors at the student level in this estimation.[22]

Table 8 shows the results of estimation (2). Interesting cases are shown in panels B and C, where Venezuelan and Haitian shocks are analyzed separately. Panel B shows that, after the Venezuelan student migratory shock, the likelihood of students transferring to a private school increased slightly (0.6%); however, this coefficient is not robust to the inclusion of controls (column 2). Likewise, higher SES students were not observed to have transferred from public to private schools.

Lastly, Panel C shows two interesting effects. The DD coefficient suggests that, after the non-Spanish-speaking migrant shock, the likelihood of treated students transferring to private schools went down by 1.3%, but higher SES students became 3.1% more likely to leave public schools for private ones. All this refers to the transition between 4th and 6th grade. It should be noted that this effect is robust to the placebo test and that the effects are observed when we include all the controls, which constitutes our preferred estimate. In addition, joint hypothesis test F, with 99% confidence, rejects the null hypothesis that that both coefficients are null, in contrast with the previous cases.

[19] In prior estimates, we only worked with students who had remained in the same school. In this section, we consider all students; otherwise, it would be impossible to estimate transfers to other schools.

[20] We follow the definition of dummy variables advanced by Farre et al. (2018). Basically, this variable indicates whether the school charges a tuition fee, which applies both to subsidized private and wholly private schools, even though the former tend to charge less than the latter. This is also observable in the recent migratory Chilean context, as 3 showed that up to two thirds of migrant students enroll in public schools; therefore, it is more consistent to seek the effects of native-flight there.

[21] A monthly income equal to or higher than $1,600,000 CLP earned by all the people living in the student's household.

[22] This estimation uses a linear probability model, following Farre et al. (2018).

Overall, since the DDD coefficient is nearly three times greater than the DD coefficient, there is evidence for the cream-skimming effect due to the arrival of non-Spanish-speaking students.

Table 8: Effects of a flow shock of immigrant students on native flight from public to private schools

	(1)	(2)	(3)	(4)
VARIABLES	4th to 6th		2nd to 4th (*placebo test*)	
Panel A: *All migrants shock*				
$T \times POST$	-0.004	0.002	-0.003	0.005*
	(0.003)	(0.003)	(0.003)	(0.003)
$T \times POST \times HighSES$	0.005	-0.006	-0.004	-0.014
	(0.009)	(0.009)	(0.009)	(0.009)
Observations	65,708	59,789	65,656	59,799
Number of students	32,884	32,884	32,884	32,884
F-test	2.67	0.93	2.01	3.90
p-value	0.264	0.629	0.367	0.142
Overall R-squared	0.0255	0.302	0.0238	0.317
Panel B: *Venezuelan shock*				
$T \times POST$	0.006**	-0.003	-0.000	0.005
	(0.003)	(0.003)	(0.003)	(0.003)
$T \times POST \times HighSES$	-0.006	0.005	-0.011	-0.021**
	(0.009)	(0.009)	(0.009)	(0.009)
Observations	54,787	49,865	54,733	49,896
Number of students	27,416	27,416	27,416	27,416
F-test	4.61	0.67	1.68	5.67
p-value	0.099	0.714	0.431	0.059
Overall R-squared	0.0288	0.300	0.0265	0.321
Panel C: *non-Spanish-speaking shock*				
$T \times POST$	-0.006	-0.013***	0.002	-0.007
	(0.005)	(0.005)	(0.005)	(0.005)
$T \times POST \times HighSES$	0.018	0.031**	-0.004	0.007
	(0.014)	(0.013)	(0.011)	(0.011)
Observations	21,528	19,459	21,517	19,442
Number of students	10,772	10,772	10,772	10,772
F-test	2.53	9.06	0.16	1.73
p-value	0.282	0.011	0.925	0.421
Overall R-squared	0.0289	0.335	0.0278	0.347
Controls	NO	YES	NO	YES

Note: Clustered standard errors at student level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The control variables of the estimations are students' gender, years of schooling of both parents, and household income level. We also include a categorical variable of the region where the school is located, a categorical variable for school type (public, private, subsidized) and socioeconomic background, and a rural school dummy. The F-test corresponds to the join null hypothesis of zero coefficients (DD and DDD interactions terms).

4.4 Additional robustness checks

Tables 12, 13, and 14 (see Appendix) refer to two exercises conducted to evaluate the robustness of the results found about the effect of migration on the academic performance of native students. First, we add native students' academic performance before the shock as a control variable, measured as the average standardized Mathematics and Language scores in 2016, when they were in 4th grade. This enables us to evaluate the results, conditional to students' past performance. Second, schools that received migrants from Venezuela or non-Spanish-speaking countries, but from no other countries, were regarded as part of the treatment group. Students coming from countries such as Bolivia, Colombia, and Peru, among others, belong to migrant communities whose inflow has varied minimally, unlike the Venezuelan and Haitian communities in recent years. Therefore, the enrollment of these students in Chilean schools does not constitute an exogenous migration shock.

The top panels of the tables show the main results of each of the shocks analyzed in the document to facilitate comparisons with the additional robustness exercises. Panel B adds native students' past performance, while Panel C regards as part of the treatment group students enrolled in schools exposed to a shock due to students from Venezuela or non-Spanish-speaking countries, but from no other countries. Although the DD coefficient varies in all cases, the change is rather modest. In any case, the estimates remain robust in terms of sign and significance in all robustness exercises. It should be noted that, in panel B of Table 12, when past performance is added as a control variable, the Reading DD coefficient displays a sizable increase compared to the rest of the panels, especially for boys. However, in panel B of table 13 (Venezuelan shock), the coefficient decreases relative to the rest of the panels, while in table 14 (Haitian shock), the coefficient increases. In both cases, the coefficients do not vary considerably in terms of magnitude; therefore, the largest increase in panel B of table 12 is offsetting capturing the changes opposite to the direction of tables 13 and 14.

5 Conclusions

This paper examines the effects of a mass immigrant influx on the educational outcomes of native students, based on the large increase in Venezuelan and Haitian migration to Chile that started in 2016 and peaked in 2018. In general, negative effects of these migratory inflows are observed. The mass arrival of immigrants, together with the relatively small size of Chile's population, made it possible to detect these effects.

To do so, the paper uses longitudinal data from administrative databases provided by the Chilean Ministry of Education and the Education Quality Agency. The identification strategy used is a differences-in-differences model in which the treated subjects are native students exposed to the migrant shock, according to the increase in the number of for-

eign students in classrooms. Most migrants attend public schools, which are free and do not conduct student selection processes. In addition, these schools concentrate lower SES native students whose parents have a lower educational level and are more economically disadvantaged in general.

The negative effects of migration are comparatively modest in Mathematics (0.04SD - 0.05SD) and are observed in both boys and girls. As for Reading, negative effects are larger (0.07SD - 0.08SD), but are only significant for boys. In addition, an interesting result emerges when considering the specific inflow of Haitian students: the negative effect on male native students' Reading performance increases by over 50% relative to the Venezuelan student shock. No negative effects are found in boys' Mathematics performance or in girls' performance in both tests. These findings can be ascribed to the fact that this population group comes from a non-Spanish-speaking country.

Unlike developed countries, Chile lacks educational programs in a foreign language. Teacher education and the resources needed to educate students in a context of diversity and complexity are far from meeting international standards. Therefore, it is relevant to take these aspects into account when considering the large flow of migrants within Latin America and the Caribbean. To the best to our knowledge, LAC school systems are not prepared for immigration and and do not possess the necessary resources to fulfill these new tasks.

Furthermore, there are no subsidies for migrant students in Chile. That is, if serving migrant students requires additional resources for meeting their cultural, language, remedial, or integration needs, a funding mechanism must be considered which will facilitate the inclusion and education of these students while reducing any potential negative effects on native-born students' learning outcomes.

Lastly, the evidence found also points to a native flight effect associated with the migratory process, as higher SES native students tend to transfer from public schools to private or subsidized ones as a result of the immigration shock. The evidence indicates that this effect is only observed after the arrival of Haitian students but not Venezuelan ones, which may be ascribed to native families' discriminatory views on this type of immigration. In consequence, it is also necessary to implement educational policies for native families to take part in cultural, racial, and linguistic integration processes.

References

Conger, D. (2015). Foreign-born peers and academic performance. Demography, 52(2), 569-592.

Contreras, D. and Gallardo, S. (2020), The effects of mass migration on natives' wages. Evidence from Chile. Working paper. Universidad de Chile.

Contreras, D., M. Brante, S. Espinoza and I. Zuñiga (2020) The effect of the Integration of Students with Special Educational Needs: Evidence from Chile. International Journal of Educational Development. Vol 74.

Diette, T. M., Uwaifo Oyelere, R. (2012). Do significant immigrant inflows create negative education impacts? Lessons from the North Carolina public school system.

Diette, T. M., Uwaifo Oyelere, R. (2014). Gender and race heterogeneity: The impact of students with limited english on native students' performance. American Economic Review, 104(5), 412-17.

Dustmann, C., Schönberg, U., Stuhler, J. (2016). The impact of immigration: Why do studies reach such different results?. Journal of Economic Perspectives, 30(4), 31-56.

Figlio, D., Özek, U. (2019). Unwelcome guests? The effects of refugees on the educational outcomes of incumbent students. Journal of Labor Economics, 37(4), 1061-1096.

Jensen, P., Rasmussen, A. W. (2011). The effect of immigrant concentration in schools on native and immigrant children's reading and math skills. Economics of Education-https://www.overleaf.com/project/5df55ecb8262e4000190fe36 Review, 30(6), 1503-1515.

Longhi, S. , Nijkamp, P. and Poot, J. (2005), A Meta-Analytic Assessment of the Effect of Immigration on Wages. Journal of Economic Surveys, 19: 451-477.

Navarro-Palau (2017) Effects of differentiated school vouchers: Evidence from a policy change and date of birth cutoffs. Economics of Education Review.Vol 58.

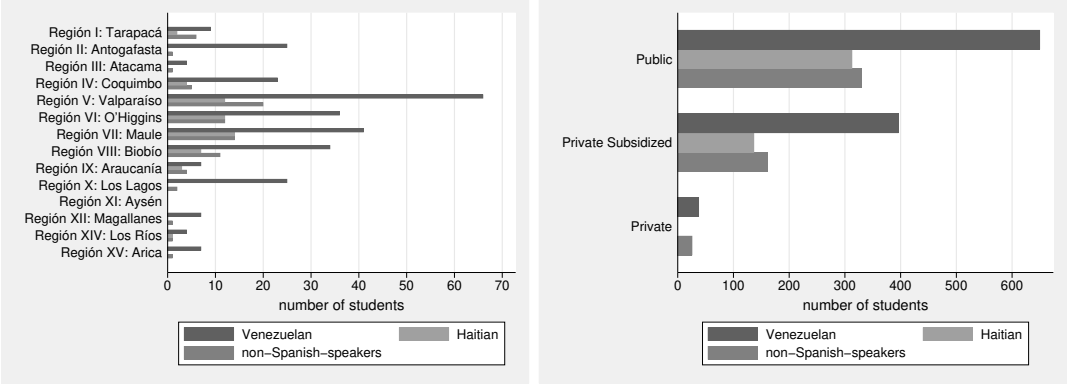
Neilson, C. A. (2013): "Targeted Vouchers, Competition Among Schools, and the Academic Achievement of Poor Students," Job market paper, Yale University.

Okkerse, L. (2008). How to measure labour market effects of immigration: A review. Journal of Economic Surveys, 22(1), 1-30

United Nations, Department of Economic and Social Affairs, Population Division (2017). International Migration Report 2017: Highlights.

Appendix

Figure 3: Distribution of Venezuelan, Haitian, and non-Spanish-speaking 6th grade students by region and school type.



(a) By region

(b) By school type

Source of migrant student data: Sistema General de Estudiantes (SIGE), 2018.

Table 9: Balance test of mean difference in covariates between controls and treated groups (Full migratory shock).

	Controls	Treated	Difference	T	p-value
Girl = 1	0.512	0.513	0.001	0.29	0.7756
Father Education (years)	11.868	11.891	0.023	0.69	0.4894
Mother Education (Years)	12.065	12.078	0.013	0.44	0.6632
Rural = 1	0.033	0.032	-0.001	0.36	0.7151
Housing income levels					
<$100.000	0.032	0.030	-0.002	1.25	0.2118
$100,000-$200,000	0.122	0.120	-0.002	0.58	0.5641
$200,000-$300,000	0.217	0.222	0.005	1.20	0.2292
$300.000-$400.000	0.171	0.175	0.003	0.94	0.3491
$400.000-$500.000	0.116	0.115	-0.001	0.45	0.6527
$500.000-$600.000	0.089	0.085	-0.004	1.54	0.1235
$600.000-$800.000	0.081	0.080	-0.001	0.38	0.7061
$800.000-$1.000.000	0.054	0.052	-0.003	1.23	0.2191
$1.000.000-$1.200.000	0.029	0.030	0.000	0.11	0.9146
$1.200.000-$1.400.000	0.017	0.017	-0.000	0.05	0.9618
$1.400.000-$1.600.000	0.013	0.013	-0.000	0.08	0.9362
$1.600.000-$1.800.000	0.009	0.008	-0.000	0.41	0.6846
$1.800.000-$2.000.000	0.010	0.011	0.001	0.84	0.4013
$2.000.000-$2.200.000	0.007	0.007	0.000	0.48	0.6343
>$2.000.000	0.032	0.036	0.004	2.31	0.0210**
Region					
Tarapacá	0.017	0.018	0.002	1.32	0.1853
Antofagasta	0.027	0.027	0.000	0.06	0.9496
Atacama	0.006	0.006	0.000	0.52	0.6065
Coquimbo	0.024	0.025	0.001	0.37	0.7105
Valparaíso	0.059	0.057	-0.002	0.80	0.4239
O'Higgins	0.041	0.048	0.008	3.84	0.0001***
Maule	0.074	0.071	-0.002	0.93	0.3523
Bío-Bío	0.043	0.040	-0.003	1.69	0.0916*
Araucanía	0.024	0.023	-0.001	0.47	0.6373
Los Lagos	0.028	0.026	-0.002	1.45	0.1472
Aysen	0.000	0.000	0.000	-	-
Magallanes	0.011	0.011	-0.000	0.25	0.8027
Metropolitana	0.615	0.613	-0.002	0.36	0.7215
Los Ríos	0.010	0.011	0.001	0.58	0.5636
Arica y Parinacota	0.021	0.023	0.001	0.97	0.3327
School type					
Public	0.435	0.428	-0.008	1.66	0.0968*
Subsidized Private	0.524	0.525	0.001	0.17	0.8675
Private	0.040	0.047	0.007	3.62	0.0003***
School Socioeconomic Background					
SES = Low	0.055	0.050	-0.005	2.31	0.0210**
SES = Medium Low	0.299	0.294	-0.005	1.19	0.2354
SES = Medium	0.429	0.434	0.005	1.07	0.2856
SES = Medium High	0.172	0.172	0.000	0.01	0.9959
SES = High	0.045	0.050	0.005	2.48	0.0133**
Observation	22,262	22,219			

Source: Fourth grade SIMCE database, 2016.

Table 10: Balance test of mean difference in covariates between controls and treated groups (Venezuelan shock).

	Controls	Treated	Difference	T	p-value
Girl = 1	0.503	0.516	0.012	2.36	0.0181**
Father Education (years)	12.088	12.052	-0.036	1.03	0.3016
Mother Education (Years)	12.232	12.204	-0.029	0.89	0.3722
Rural = 1	0.025	0.022	-0.003	1.87	0.0620*
Housing income levels					
<$100.000	0.028	0.026	-0.002	1.26	0.2068
$100,000-$200,000	0.110	0.109	-0.001	0.21	0.8316
$200,000-$300,000	0.204	0.213	0.009	2.05	0.0408**
$300.000-$400.000	0.173	0.177	0.004	0.99	0.3233
$400.000-$500.000	0.119	0.118	-0.001	0.24	0.8115
$500.000-$600.000	0.092	0.089	-0.003	1.06	0.2885
$600.000-$800.000	0.087	0.087	0.000	0.00	0.9971
$800.000-$1.000.000	0.059	0.056	-0.003	1.30	0.1942
$1.000.000-$1.200.000	0.032	0.032	0.000	0.07	0.9425
$1.200.000-$1.400.000	0.019	0.018	-0.000	0.30	0.7639
$1.400.000-$1.600.000	0.015	0.015	-0.000	0.16	0.8752
$1.600.000-$1.800.000	0.009	0.009	0.000	0.00	0.9982
$1.800.000-$2.000.000	0.012	0.012	0.000	0.23	0.8180
$2.000.000-$2.200.000	0.007	0.008	0.000	0.11	0.9159
>$2.000.000	0.033	0.031	-0.003	1.36	0.1724
Region					
Tarapacá	0.017	0.018	0.001	0.83	0.4038
Antofagasta	0.032	0.032	-0.000	0.02	0.9854
Atacama	0.004	0.005	0.000	0.70	0.4836
Coquimbo	0.020	0.018	-0.001	0.96	0.3368
Valparaíso	0.055	0.053	-0.001	0.63	0.5277
O'higgins	0.037	0.044	0.007	3.46	0.0005***
Maule	0.070	0.073	0.003	1.22	0.2241
Bío-Bío	0.040	0.036	-0.004	2.05	0.0402**
Araucanía	0.021	0.019	-0.001	0.90	0.3698
Los Lagos	0.033	0.030	-0.002	1.40	0.1626
Aysen	0.000	0.000	0.000	-	-
Magallanes	0.012	0.013	0.000	0.32	0.7503
Metropolitana	0.625	0.620	-0.004	0.89	0.3741
Los Ríos	0.011	0.012	0.001	0.71	0.4761
Arica y Parinacota	0.024	0.026	0.002	1.32	0.1883
School type					
Public	0.445	0.438	-0.007	1.37	0.1714
Subsidized Private	0.513	0.520	0.006	1.23	0.2170
Private	0.042	0.043	0.001	0.31	0.7573
School Socioeconomic Background					
SES = Low	0.035	0.030	-0.005	2.57	0.0101**
SES = Medium Low	0.276	0.263	-0.012	2.70	0.0069***
SES = Medium	0.445	0.465	0.020	3.98	0.0001***
SES = Medium High	0.196	0.195	-0.001	0.26	0.7938
SES = High	0.048	0.046	-0.002	1.05	0.2946
Observations	19184	19109			

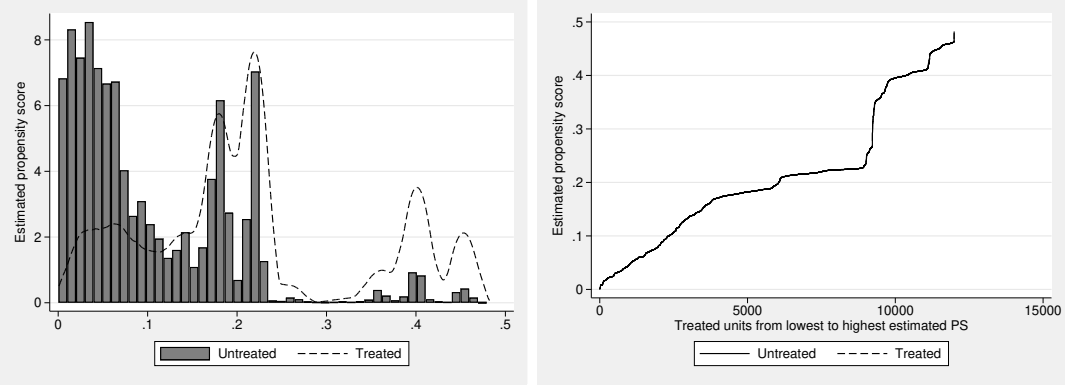
Source: Fourth grade SIMCE database, 2016.

Table 11: Balance test of mean difference in covariates between controls and treated groups (non-Spanish-speaking migrant shock).

	Controls	Treated	Difference	T	*p-value*
Girl = 1	0.523	0.526	0.003	0.26	0.7964
Father Education (years)	11.053	11.015	-0.038	0.46	0.6429
Mother Education (Years)	11.439	11.399	-0.041	0.56	0.5737
Rural = 1	0.048	0.053	0.005	0.91	0.3610
Housing income levels					
<$100.000	0.037	0.039	0.001	0.34	0.7374
$100,000-$200,000	0.126	0.124	-0.002	0.29	0.7715
$200,000-$300,000	0.247	0.243	-0.005	0.50	0.6179
$300.000-$400.000	0.194	0.193	-0.001	0.06	0.9507
$400.000-$500.000	0.133	0.131	-0.003	0.35	0.7232
$500.000-$600.000	0.087	0.086	-0.001	0.08	0.9365
$600.000-$800.000	0.062	0.065	0.003	0.48	0.6282
$800.000-$1.000.000	0.031	0.032	0.001	0.24	0.8107
$1.000.000-$1.200.000	0.018	0.021	0.002	0.73	0.4674
$1.200.000-$1.400.000	0.013	0.011	-0.002	0.95	0.3406
$1.400.000-$1.600.000	0.004	0.006	0.002	0.99	0.3204
$1.600.000-$1.800.000	0.006	0.006	-0.000	0.01	0.9923
$1.800.000-$2.000.000	0.004	0.005	0.001	0.49	0.6273
$2.000.000-$2.200.000	0.004	0.005	0.001	0.86	0.3877
>$2.000.000	0.032	0.034	0.002	0.55	0.5794
Region					
Tarapacá	0.008	0.011	0.003	1.30	0.1939
Antofagasta	0.000	0.000	0.000	-	-
Atacama	0.006	0.005	-0.001	0.32	0.7499
Coquimbo	0.024	0.028	0.004	1.13	0.2603
Valparaíso	0.045	0.043	-0.002	0.42	0.6758
O'higgins	0.031	0.038	0.007	1.67	0.0948*
Maule	0.070	0.075	0.005	0.85	0.3961
Bío-Bío	0.035	0.033	-0.003	0.66	0.5112
Araucanía	0.022	0.020	-0.002	0.58	0.5623
Los Lagos	0.003	0.002	-0.001	1.22	0.2225
Aysen	0.000	0.000	0.000	-	-
Magallanes	0.002	0.001	-0.001	1.00	0.3150
Metropolitana	0.748	0.739	-0.008	0.82	0.4127
Los Ríos	0.002	0.002	-0.001	0.51	0.6127
Arica y Parinacota	0.004	0.003	-0.001	0.79	0.4277
School type					
Public	0.466	0.474	0.009	0.78	0.4367
Subsidized Private	0.501	0.487	-0.014	1.20	0.2288
Private	0.034	0.039	0.005	1.14	0.2533
School Socioeconomic Background					
SES = Low	0.125	0.128	0.003	0.37	0.7145
SES = Medium Low	0.423	0.426	0.002	0.19	0.8466
SES = Medium	0.351	0.350	-0.001	0.12	0.9071
SES = Medium High	0.066	0.058	-0.008	1.50	0.1325
SES = High	0.034	0.039	0.005	1.08	0.2806
Observations	3791	3802			

Source: Fourth grade SIMCE database, 2016.

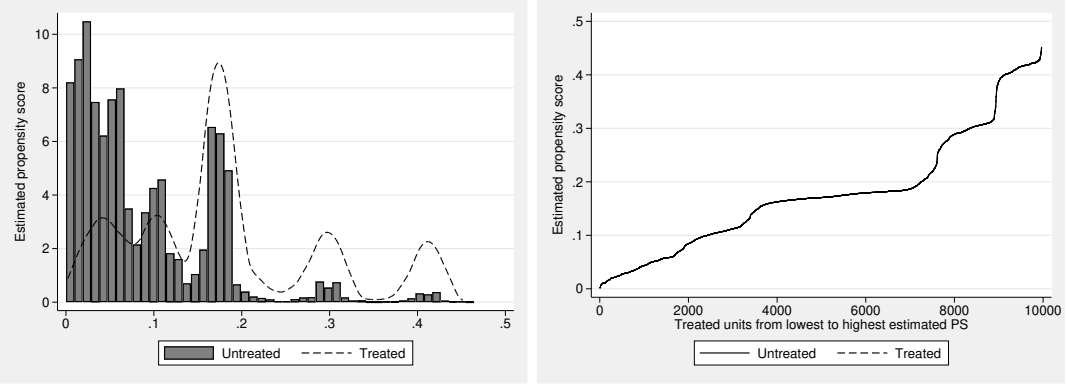
Figure 4: Propensity Score Matching (Full migratory shock)



(a) Common Support

(b) Lowest To Highest PS

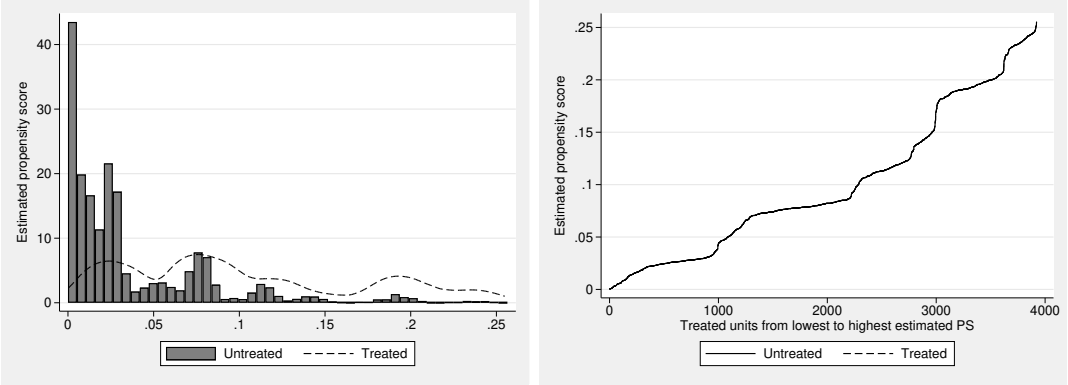
Figure 5: Propensity Score Matching (Venezuelan shock)



(a) Common Support

(b) Lowest To Highest PS

Figure 6: Propensity Score Matching (Haitian shock)



(a) Common Support

(b) Lowest To Highest PS

Table 12: Additional robustness checks: Full immigration flow shock.

	Reading			Math		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Full	Boys	Girls	Full	Boys	Girls
Panel A: *Main results*						
Difference-in-Difference	-0.045***	-0.066***	-0.028*	-0.045***	-0.051***	-0.041***
	(0.010)	(0.015)	(0.014)	(0.010)	(0.015)	(0.014)
Observations	42,020	20,509	21,511	42,221	20,581	21,640
Number of students	23,563	11,568	12,012	23,611	11,581	12,046
Overall R-squared	0.102	0.0980	0.0992	0.142	0.148	0.135
Panel B: *Adding past performance as a control variable*						
DD	-0.054***	-0.080***	-0.030**	-0.051***	-0.060***	-0.045***
	(0.010)	(0.015)	(0.014)	(0.010)	(0.014)	(0.014)
Observations	41,586	20,249	21,337	41,655	20,266	21,389
Number of students	23,198	11,353	11,861	23,198	11,353	11,861
Overall R-squared	0.702	0.702	0.700	0.722	0.723	0.719
Panel C: *Excluding other immigration shock*						
DD	-0.047***	-0.059***	-0.036**	-0.057***	-0.060***	-0.055***
	(0.012)	(0.017)	(0.016)	(0.011)	(0.016)	(0.015)
Observations	35,042	17,271	17,771	35,211	17,330	17,881
Number of students	19,594	9,713	9,897	19,628	9,721	9,922
Overall R-squared	0.105	0.101	0.103	0.147	0.151	0.141

Note: Clustered standard errors at student level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The sample is limited to native students who did not transfer schools between 2014 (when they attended 2nd grade) and 2018 (when they attended 6th grade). The control variables of the estimations are gender, years of schooling of both parents, and the student's household income level. We also include a categorical variable for the region where the school is located, a categorical variable for school type (public, private, subsidized) and socioeconomic background, and a rural school dummy. Panel B, includes past performance as a control variable (average of Mathematics and Language test score of 4th grade students). Panel C excludes schools exposed to an inflow of migrant students (other than Venezuelans and Haitians) for the first time.

Table 13: Additional robustness checks: Venezuelan student flow shock.

	Reading			Math		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Full	Boys	Girls	Full	Boys	Girls
Panel A: *Main results*						
Difference-in-Difference	-0.037***	-0.058***	-0.019	-0.053***	-0.054***	-0.053***
	(0.011)	(0.017)	(0.016)	(0.011)	(0.016)	(0.015)
Observations	35,009	17,128	17,881	35,161	17,177	17,984
Number of students	19,611	9,641	9,983	19,647	9,649	10,010
Overall R-squared	0.101	0.0964	0.101	0.142	0.144	0.137
Panel B: *Adding past performance as a control variable*						
Difference-in-Difference	-0.025**	-0.054***	-0.000	-0.047***	-0.050***	-0.044***
	(0.011)	(0.017)	(0.016)	(0.011)	(0.016)	(0.015)
Observations	34,672	16,935	17,737	34,729	16,952	17,777
Number of students	19,331	9,484	9,859	19,331	9,484	9,859
Overall R-squared	0.692	0.692	0.692	0.721	0.721	0.719
Panel C: *Excluding other immigration shock*						
Difference-in-Difference	-0.039***	-0.057***	-0.024	-0.064***	-0.064***	-0.065***
	(0.013)	(0.018)	(0.017)	(0.012)	(0.017)	(0.017)
Observations	29,215	14,494	14,721	29,333	14,526	14,807
Number of students	16,329	8,139	8,202	16,352	8,142	8,221
Overall R-squared	0.106	0.104	0.104	0.148	0.152	0.142

Note: Clustered standard errors at student level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Same details as Table 12.

Table 14: Additional robustness checks: non-Spanish-speaking student flow shock.

	Reading			Math		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Full	Boys	Girls	Full	Boys	Girls
Panel A: *Main results*						
Difference-in-Difference	-0.055***	-0.084***	-0.028	-0.020	-0.020	-0.021
	(0.018)	(0.028)	(0.025)	(0.018)	(0.026)	(0.024)
Observations	13,524	6,375	7,149	13,592	6,388	7,204
Number of students	7,689	3,656	4,039	7,708	3,660	4,054
Overall R-squared	0.111	0.110	0.103	0.151	0.161	0.143
Panel B: *Adding past performance as a control variable*						
Difference-in-Difference	-0.061***	-0.098***	-0.029	-0.025	-0.027	-0.024
	(0.018)	(0.028)	(0.025)	(0.018)	(0.026)	(0.024)
Observations	13,372	6,290	7,082	13,390	6,293	7,097
Number of students	7,559	3,588	3,977	7,559	3,588	3,977
Overall R-squared	0.707	0.706	0.705	0.722	0.732	0.713
Panel C: *Excluding other immigration shock*						
Difference-in-Difference	-0.055**	-0.071**	-0.042	-0.024	-0.026	-0.023
	(0.021)	(0.033)	(0.028)	(0.020)	(0.030)	(0.028)
Observations	10,457	4,934	5,523	10,507	4,943	5,564
Number of students	5,929	2,818	3,117	5,946	2,822	3,130
Overall R-squared	0.115	0.112	0.110	0.158	0.168	0.152

Note: Clustered standard errors at student level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Same details as Table 12.