CEDLAS
FCE FACULTAD DE CIENCIAS ECONÓMICAS
UNIVERSIDAD NACIONAL DE LA PLATA

DOCUMENTOS DE TRABAJO

Effects of Immigration on Native Learning: The Case of the Venezuelan Crisis

Nicolás Irazoque Sillerico

Documento de Trabajo Nro. 364
Enero, 2026
ISSN 1853-0168
www.cedlas.econo.unlp.edu.ar

Cita sugerida: Irazoque Sillerico, N. (2026). Effects of Immigration on Native Learning: The Case of the Venezuelan Crisis. Documentos de Trabajo del CEDLAS Nº 364, Enero, 2026, CEDLAS-Universidad Nacional de La Plata.

Effects of Immigration on Native Learning: The Case of the Venezuelan Crisis

Nicolás Irazoque Sillerico *

August 16, 2025

[Download the latest version]

Abstract

This paper provides the first estimate of the impact of the Venezuelan exodus on Colombian students' learning. To identify the impact, I use the reopening of the Colombian-Venezuelan border in 2016 as a natural experiment and propose a differences-in-differences design. The results indicate that, on average, native high school students exposed to immigrants on the schools experience a decrease of 1.8% of a standard deviation in their academic performance and the effect is persistent for the first four years and tends to zero after that. A possible mechanism for this negative effect is that teachers allocate class time to assist lower-achieving Venezuelans. This effect becomes insignificant when the concentration of immigrants is higher. The negative effect is larger for women, for Colombians with high achievement, with highly educated mothers, and for natives who attend schools with high average scores and a high concentration of educated mothers.

JEL Classification: I21, J15, J24.

Key words: Immigration, Education, Immigrant children, Peer effects.

1 Introduction

In recent years, there has been an increase in migratory flows around the world, leading to discussions about their impacts on both the host population and the migrants. One of the most significant examples is the *"Venezuelan exodus,"* where millions of Venezuelans have fled their, primarily to Colombia, due to political, social, and economic crises. Between 2014 and 2020, approximately 4.6 million Venezuelans emigrated, with 1.8 million settling in Colombia (UNHCR, 2021). Estimates by the National Administrative Department of Statistics (DANE, for its abbreviation in Spanish) indicated that the Venezuelan population accounted for approximately 4.4% of Colombia's total population by the end of 2020 (DANE, 2021).

While there is existing literature focused on the impact on the Colombian labor market (Bahar et al., 2021; Caruso et al., 2019; Delgado-Prieto, 2021; Lebow, 2021; Pedrazzi & Peñaloza-Pacheco, 2020; Santamaria, 2020), an unexplored area concerns the educational outcomes, specifically how they may be affected for native Colombian students. The influence of

*Independent Researcher. Contact: nico.irazoque@gmail.com. I thank Florencia Pinto (CEDLAS), Matías Busso (IDB), Julian Cristia (IDB), Natalie Echenique (Aru Foundation), Ricardo Nogales (OPHI) and Ernesto Yañez (UNESCO) for helpful comments and discussions. This work is an improved and extended version of my Masters' Thesis at UNLP. Any errors in the paper are my responsibility.

Venezuelan immigrants on these outcomes remains unclear. On one hand, the *Peer Average Background Effect* (Figlio et al., 2021; van der Werf, 2021) suggests that interactions with immigrant peers from high socioeconomic backgrounds could benefit native students. Additionally, the *Peer Group Composition Effect* (Figlio & Özek, 2019) proposes that segregating low-performing immigrants and native students into separate classrooms might also prove advantageous for natives. On the other hand, if immigrant students require more teacher attention due to lower proficiency in the language of instruction (*Peer Ethnic Composition Effect*) or their low socioeconomic and educational backgrounds, this could lead to negative peer effects that deteriorate native students' learning outcomes (Ballatore et al., 2018; Gould et al., 2009; Jensen & Rasmussen, 2011; Tonello, 2016). However, some studies report no significant impact (Bossavie, 2020; Contini, 2013; Frattini & Meschi, 2019; Ohinata & Van Ours, 2013).

The literature attributes positive effects primarily to two factors: the *Peer Group Composition Effect* and the *Peer Average Background Effect*. Figlio and Özek (2019) present results supporting the first factor, showing that low-achieving immigrant students concentrated in separate classrooms from native students yield beneficial outcomes.[1] Figlio et al. (2021) and van der Werf (2021) provide evidence for the second factor, grouping native-born students with low-education parents and immigrant students from higher socioeconomic backgrounds.[2] Conversely, negative effects are often linked to the *Peer Ethnic Composition Effect* or to the inferior quality of schools that immigrants attend. This factor is particularly significant when immigrants' native language differs substantially from the instruction language, leading to lower academic performance for native students (Ballatore et al., 2018; Tonello, 2016; Jensen & Rasmussen, 2011).[3] This mechanism appears relevant in studies that identify non-significant effects.[4] Lastly, Gould et al. (2009) indicates that higher immigrant concentrations may adversely affect the likelihood of passing high school exams, exacerbated by the lower educational quality of schools accommodating immigrants.[5]

The aim of this paper is to analyze the impact of Venezuelan immigration on the learning outcomes of Colombian students. For this purpose, I utilize the mid-2016 reopening of the Colombian-Venezuelan border as a natural experiment. The primary data sources include the Saber 11 test scores of eleventh-grade students between 16 and 17 years old.[6] The identification strategy exploits the variation in the number of Venezuelan migrants across schools in Colombia following the border's reopening. To estimate the causal effect on the academic performance of Colombian students from 2013 and 2022, I propose to utilize a difference-in-differences design. I define the treatment group as schools that received at least one Venezuelan immigrant after the border reopened that are taking the Saber 11 tests, and the comparison group as schools that did not had any immigrants taking the tests.

The results show that the academic performance of native Colombians in the treated schools decreased by an average of 1.8% of one standard deviation compared to schools without immigrants taking the tests. The results are consistent with the inclusion of various controls and show an effect size ranging from 0.008 to 0.024 standard deviations across different outcomes,

[1] Figlio and Özek (2019) analyzed the immigration of Haitian people to Florida, United States.

[2] van der Werf (2021) examined Vietnamese immigration to the United States following the Vietnam War, with at least one college-educated parent among the immigrant families.

[3] Ballatore et al. (2018) and Tonello (2016) studied immigrant groups in Italy unfamiliar with Italian, while Jensen and Rasmussen (2011) focused on Denmark.

[4] Research by Bossavie (2020) and Ohinata and Van Ours (2013) in the Netherlands, Contini (2013) and Frattini and Meschi (2019) in Italy, and Green et al. (2022) in Norway illustrates these dynamics.

[5] Gould et al. (2009) investigated the migration impacts from the former Soviet Union to Israel during the 1990s.

[6] I also use the Great Integrated Household Survey (GEIH, for its abbreviation in Spanish) Migration Module to identify migration patterns for an extension of the results.

such as math, reading, science and social science, and English. This effect corresponds to about 12% to 8.6% of what a student typically learns in an academic year. Event study results suggests that the negative effect is persistent since the first year of immigrants arrival up to the fourth year and it gets smaller and disappears from the fifth year. The effects are somewhat robust to different estimates with sample modifications.

Heterogeneities for different groups suggests large variations between groups. I found differences between men and women, where the negative effect is larger for women. Based on the education level of mothers, I found that the negative effect is consistently higher for students with highly educated mothers. For natives with high performance the effect is always negative and for low-performing natives the effect is not-significant for natural and social sciences, and the summary score. The negative effect is higher for students attending schools with a high concentration of educated mothers and schools with a high average score, than the effect for their counterparts. In general, effects are not-significant for students in schools with low concentration of educated mothers and low-performing schools.

One potential mechanism is teacher allocation of class time. As Venezuelans tend to have lower academic performance, compared to Colombians, and intensive effects suggests that an increase of 1% of immigrants in treated schools increases natives learning 1.2%, I believe that teachers allocate more class time to help them, causing an indirect negative peer effect over native students when the share of immigrants is low. as soon as the the share of immigrants is higher, the negative effect disappears. This suggests that teachers adapt its pedagogical practices or schools hires new professors when the share of low performing immigrants increases. I also found that this mechanism could disappear in schools where natives and immigrants tend to have similar performance and the effect could turn positive if the share of immigrants by departments is large.

The contributions to the literature on migration and education of this paper can be summarized as follows. Firstly, this paper extends the literature examining the impact of immigration on the learning of native students in middle income level countries. To the best of my knowledge, the only studies of middle income countries Martínez and Martínez (2023) and Tumen (2021) analyses of Peru and Turkey, respectively. Appendix A Table A.1 presents all of the identified papers comparing income level and GDP per capita, showing Peru as the country with the lowest income level of destination countries for this literature with 12,553.8 PPP[7] in 2022. I contribute by expanding the knowledge of immigrant effects on natives learning in middle income countries focusing in Colombia, which had a GDP per capita of 14,661 PPP in 2022.

Secondly, I contribute to the existing literature on the impact of large-scale immigration on native learning following crises, such as natural disasters, wars, or economic crisis (Figlio & Özek, 2019; Gould et al., 2009; Tumen, 2021; van der Werf, 2021, to name a few examples). Finally, I also provide specific findings on the phenomenon of Venezuela's mass immigration. Previous research has examined its impact on the Colombian labor market (Bahar et al., 2021; Caruso et al., 2019; Delgado-Prieto, 2021; Lebow, 2021; Peñaloza-Pacheco, 2019; Santamaria, 2020, are some examples), trade (Lombardo & Peñaloza-Pacheco, 2021), and consumption levels (Ibánez et al., 2021). Other papers such as Contreras and Gallardo (2022), Luksic (2022) and Martínez and Martínez (2023) studied the effects on the Chilean and Peruvian education system, but the migratory flow is small compared to the Colombian case. This paper adds to this literature by analyzing its impact on the education sector in Colombia, which received the highest flow of Venezuelans.

The paper is organized as follows. Section 2 provides an overview of the Venezuelan crisis and immigration to Colombia. Section 3 presents the data sources and descriptive statistics.

[7]Constant 2017 international $us, World Bank estimates.

Section 4 explains the identification strategy. The results for natives, robustness checks and heterogeneous effects are presented in Section 5. Section 6 explores a potential mechanism and some extensions to the results. I discuss some limitations and next steps for the research in Section 7. Finally, the conclusions are discussed in Section 8.

2 The crisis and the Venezuelan exodus

The political and economic crisis in Venezuela began in late 1998 following the election of President Hugo Chávez, whose administration was marked by nationalizations, restrictions on the private sector, expropriations, and the implementation of large-scale social programs funded with resources obtained from the international oil boom of the 2000s (Bahar et al., 2021). Chávez's presidency ended with his death in 2013, leaving Venezuela with a weakened private sector, a declining oil industry (Hernandez & Monaldi, 2016), and a tense diplomatic relations with neighboring countries (Romero, 2008). Following the presidential elections in April 2013, his successor Nicolas Maduro continued the same policies of the Chávez regime. Under Maduro's leadership, instances of authoritarianism and abuse of power have increased, and the Venezuelan economy has shown signs of imminent crisis.

Tensions between Colombia and Venezuela escalated in 2015 due to allegations of Colombian armed groups within Venezuelan borders. Maduro's declaration of a state of emergency and the subsequent closure of the border in Táchira led to halted migratory flows and trade, and the deportation of over 180 Colombian citizens. These actions intensified diplomatic strains and culminated in the complete shutdown of the Colombian-Venezuelan borders in August 2015.

In 2016, following massive protests in Venezuela, the borders were temporarily opened in July and fully reopened in August. The economic and social crisis in Venezuela, characterized by hyperinflation, lack of access to basic services, and increased violence in the country, led to a massive exodus of Venezuelan citizens to neighboring Colombia in the second half of 2016 (Santamaria, 2020).[8]

According to the United Nations High Commissioner for Refugees (UNHCR), over 5.2 million Venezuelans have left their country due to ongoing crises, with Colombia being the primary destination. Other significant host countries include Argentina, Brazil, Chile, Peru, and Ecuador. Colombian government data projected that by 2021, approximately 1.8 million Venezuelans would reside within its borders, predominantly entering through the Norte de Santander department, a key transit point.

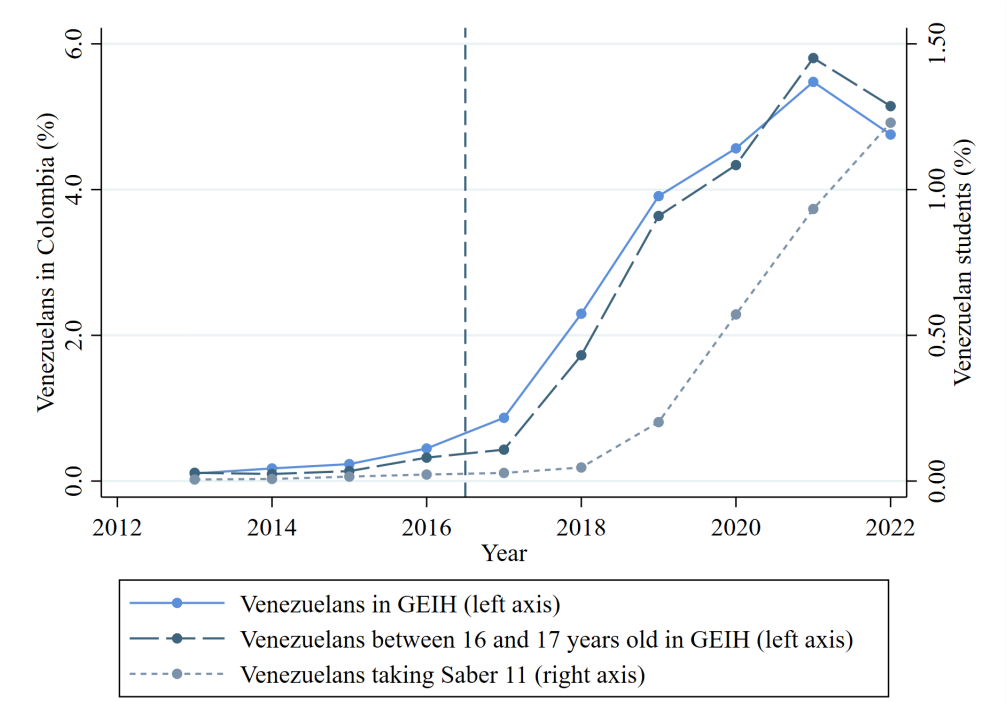
In response to this influx, the Colombian government instituted the "Special Permit of Permanence" (PEP, for its abbreviation in Spanish), a residency permit aimed at integrating Venezuelans by granting them access to the labor market and state services like education and healthcare. For educational integration, students can validate their Venezuelan school records with the Ministry of Education to be placed in the appropriate grade. If records are unavailable, students are administered knowledge tests and assigned to the grade level based on the test results (Ministerio de Educación, 2015). A timeline of the above events is presented in Table A.2 in the Appendix A.

To visualize this migratory exodus, I present Figure 1, which shows the proportion of Venezuelan immigrants in Colombia based on the GEIH and the Saber 11 tests. Using the

[8]By 2018, Venezuela's situation had escalated to a humanitarian crisis. Reports indicated that between 87% and % of households lived in poverty, 75% of the population had lost an average of 11 kilograms due to food shortages, and 33% ate no more than twice a day (Sequera, 2018). Additionally, infant mortality rates increased by a minimum of 30% and maternal mortality rates reached 65% (The Guardian, 2017). The controversial re-election of Maduro in mid-2018 may have contributed to a peak in the exodus of Venezuelans to Colombia (Santamaria, 2020).

GEIH Migration Module, I compute the number of Venezuelans as a proportion of the population of Colombia, as well as the proportion of Venezuelans among the population aged 16 and 17 (relevant ages for this study). The Saber 11 Tests allow calculating the proportion of Venezuelans among students in 11th grade. The analysis of both data sources indicates a significant increase in the proportion of Venezuelan immigrants in Colombia from 2015. Between 2015 to 2021, the proportion of Venezuelans in the total population increased from 0.2% to 5.5%, a jump of 5.3 percentage points. Between 2015 and 2022, the proportion of Venezuelans among Colombia's 11th graders increased from 0.01% to 1.2%, a 120 times increase. Similarly, looking at the population between 16 and 17 years old, the proportion of Venezuelans increased by more than 40 times, from 0.1% in 2015 to 5.8% in 2020.

Figure 1: Migratory flows of Venezuelans to Colombia (2012-2022)



Note: The light blue line, with a solid line, and the dark blue line, with long dashes, both measured on the left axis, correspond to the proportion of all Venezuelans and Venezuelans between 16 and 17 years old in Colombia, respectively. The light blue line, with short dash, measured on the right axis, corresponds to the proportion of Venezuelan students taking the Saber 11 tests each year. The sampling weights from the surveys were used to calculate the annual proportions. The dark blue vertical line, with dashes, corresponds to the opening of borders between Colombia and Venezuela in mid-2016.
Source: Own elaboration using Saber 11 tests from 2012 to 2022 and GEIH Migration Module from 2013 to 2022.

3 Data sources and descriptive statistics

3.1 Student learning

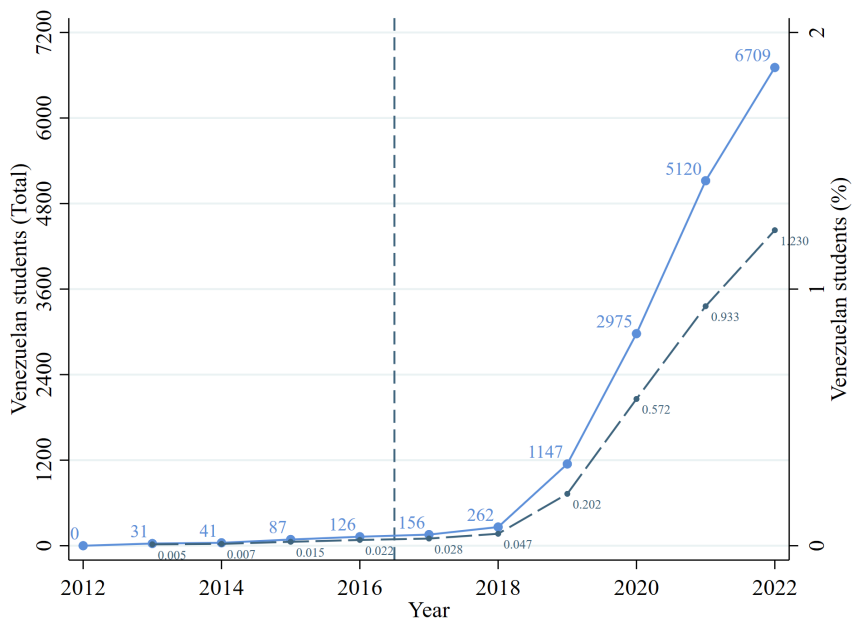
Information on the learning outcomes of Colombian students is collected through the Saber 11 tests, administered by the Colombian Institute for the Evaluation of Education (ICFES, for its abbreviation in Spanish). The Saber 11 tests are standardized tests that provide information on the competencies of students at the end of secondary education (middle school in Colombia). The tests are administered twice a year, in the first semester for schools that follow the B

calendar (September to June) and in the second semester for schools that follow the A calendar (February to November).[9]

Since 2000, the tests' structure has undergone modifications, but as of the second semester of 2014, the structure enables comparability between years. The similarity of the test structures from year to year facilitates the comparison across years. The competencies evaluated include critical reading, mathematics, natural sciences, social science, and English. [10]

The tests show a notable increase in the number of Venezuelan students taking the exam between 2012 and 2022. Figure 2 shows the total number and percentage of Venezuelan students who took the tests between these years. Two main things can be observed. First, the number of Venezuelan students increased more than 200 times between 2013 and 2022, from 31 Venezuelans in 2013 to over 6,700 in 2022. This means that the relative increase of Venezuelans in the total universe of students went from 0.005% to 1.23% between the two years.[11] Second, the increase was more pronounced after the border reopened in 2016. From 2013 to 2016, the number of Venezuelan test-takers nearly quintupled, and from 2016 to 2022, it increased more than 55 times.

Figure 2: Venezuelan students taking Saber 11 tests (2012-2022)



Note: The light blue line, with a solid line, measured on the left axis corresponds to the total number of Venezuelan students taking Saber 11 tests each year. The dark blue line, with long dashes, measured on the right axis, corresponds too the percentage of students taking the Saber 11 tests each year. The dark blue vertical line, with dashes, corresponds to the opening of borders between Colombia and Venezuela in mid-2016.
Source: Own elaboration using Saber 11 tests from 2012 to 2022.

On the other hand, Figure A.1 in the Appendix A, shows the increase of schools with at least one Venezuelan student in their 11th grade classrooms. Between 2013 and 2022, the number of schools with at least one Venezuelan student increased by about 128 times, from 24 schools in 2013 to 3,090 in 2022. This represents a relative increase from 0.2% to over 20% of schools

[9]On average, only 2.8% of schools in Colombia follow the B calendar.

[10]Due to a methodological change in the tests in 2014, ICFES implemented re-qualifications since the second test in 2012 in order to add periods that are comparable with the new structure.

[11]The total number of students is between 540,000 and 550,000.

with Venezuelans between these years.[12]

3.2 Geographical location of immigrants

Identifying the new location of immigrants after a large migration flow is a challenge. The primary data source I use is the GEIH provided by DANE. The survey provides comprehensive information on individuals in Colombia, including their age, sex, level of education, municipality or department of residence, type of work, and income levels. The surveys contain representative data from 23 departments, 13 metropolitan areas, and 11 intermediate cities in Colombia[13] and have been conducted monthly since 2006.

The GEIH Migration Module collects data on place of birth, places of residence over the past year and the preceding five years, and reasons for residence changes. Although the Migration Module began in 2012; however, for this paper, I use the information provided between 2013 and 2019, as questions on place of birth were not asked in the first year.

Figure 3 provides a visual representation of the proportions of Venezuelan immigrants in Colombia, by department, for the years 2013 and 2019. Data was gathered from individuals who reported in the migration module that they were born in Venezuela. The first finding is that in practically all Colombian departments there has been a considerable increase in the number of Venezuelans residing there. The relative increase of Venezuelan immigrants in Colombia is primarily concentrated in border departments. La Guajira (13.8%), Norte de Santander (11.4%), and Magdalena (10.8%) were among the border departments with the highest proportion of Venezuelans in 2019.

3.3 Characteristics of Venezuelan immigrants

Table 1 provides an overview of the composition of Venezuelan immigrant students and Colombian natives, focusing on their test scores and demographic and economic characteristics. It is observed that Venezuelan immigrants, on average, achieve lower test scores than their Colombian counterparts. The most pronounced disparity is in Mathematics, where immigrant scores are 0.18 standard deviations below those of Colombians. For other subjects, the negative differential for immigrants ranges from 0.06 to 0.11 standard deviations. On average, natives exhibit a test score advantage of 0.12 standard deviations.

Approximately 54% of Colombian students are female, compared to 59% of Venezuelan students. However, only a small proportion of students reside in rural areas, with 15% being natives and 9% immigrants, suggesting that the majority of Venezuelan students have migrated to urban areas. Furthermore, it is observed that Venezuelan students tend to live in households within higher socio-economic strata compared to Colombian students.[14] Specifically, a larger proportion of Colombians are categorized under strata 1 and 2, whereas strata 4, 5, and 6 see higher representation among Venezuelans, indicating that, on average, immigrant families are comparatively wealthier than Colombian families.

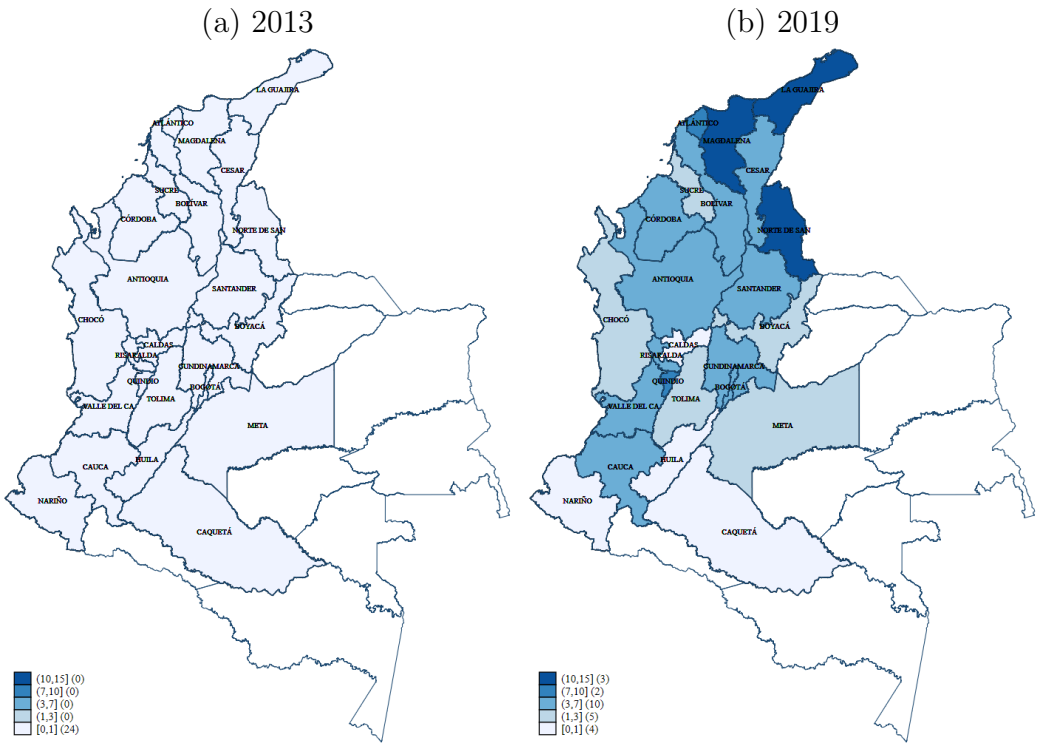
Notably, the Saber 11 tests allow for an analysis of the students' parental educational backgrounds. It is observed that 45% of Colombians have mothers with low education (i.e. whose highest level of education is incomplete secondary education), versus 32% of Venezuelan

[12]The total number of schools presenting students for the tests annually ranges between 11,500 and 13,500.

[13]Colombia has a total of 32 departments. The 9 departments that are not covered by the GEIH represent 3% of the students taking the Saber 11 test and 3% of the Colombian population according to the 2005 census.

[14]The Colombian stratification system, developed by Law 142 of 1994, functions as a cross-subsidy mechanism where users with greater payment capacity subsidize those with less (Eslava et al., 2021). The system provides subsidies to users in strata 1, 2, and 3, funded by those in strata 5 and 6, and commercial users.

Figure 3: Venezuelan immigrants by department (%)



Note: Venezuelan immigrants are considered those individuals who reported being born in Venezuela in the responses to the Migration Module. The departments with white color, which do not have the corresponding name, do not have observations in the GEIH. Most of the departments without observations area located in the Amazon region and have a low population density. The sampling weights from the survey carried out in the month of August were used to calculate the departmental proportions.
Source: Own elaboration using GEIH Migration Module of 2013 and 2019.

students. In contrast, 33% of natives versus 34% of immigrants have mothers with medium education (complete secondary education or incomplete higher education), and 22% compared to 33% have mothers with high education (complete higher education or more), respectively. The education levels of fathers show similar patterns, with a higher proportion of both groups having low education levels. Overall, the parents of Venezuelan students exhibit higher educational attainment than their Colombian counterparts.

On the other hand, Table A.3 in Appendix A presents the same descriptive statistics at the school level, considering the average per school for natives only. It differentiates the sample into schools attended by immigrants (4,950 schools) and those exclusively for natives (15,152 schools). The data reveal that schools with Venezuelan immigrants generally have higher average achievements and students from higher socioeconomic backgrounds, as indicated by parental education levels and economic strata.

To explore the potential reasons for the seemingly superior educational and socio-demographic profiles of Venezuelan immigrants' parents, I refer to Table A.2 in Appendix A. This table presents the mean scores achieved by Colombia and Venezuela in the First Regional Comparative and Explanatory Study (PERCE, as abbreviated in Spanish) of 1997, conducted by UNESCO.[15] In this study, standardized tests in mathematics and reading were administered to third and fourth-grade students. Colombia displayed higher average performances in mathematics across both grades. In reading, Colombia showed a slight advantage at the fourth grade level, while Venezuela had a marginal lead in the third grade, although this difference is not

[15]It is important to note that this international test took place more than a year before the first presidential election in which Hugo Chávez was elected president of Venezuela.

Table 1: Characteristics of Venezuelan immigrants and Colombian natives (2013-2022)

Variable	Natives		Immigrants		
	Average	Standard Dev	Average	Standard Dev	Difference
Summary	0.04	1.00	−0.08	0.96	0.12***
Math	0.04	0.99	−0.14	0.97	0.18***
Reading	0.04	0.99	−0.03	0.95	0.07***
Natural Sciences	0.04	0.99	−0.07	0.95	0.11***
Social Sciences	0.03	1.00	−0.06	0.94	0.09***
English	0.03	1.00	−0.03	1.04	0.06***
Female (%)	0.54	0.50	0.59	0.49	−0.05***
Age	17.21	2.18	17.68	2.27	−0.47***
Rural school (%)	0.15	0.35	0.09	0.29	0.05***
Low mother education (%)	0.45	0.50	0.32	0.47	0.13***
Middle mother education (%)	0.33	0.47	0.34	0.47	−0.02***
High mother education (%)	0.22	0.42	0.33	0.47	−0.11***
Low father education (%)	0.52	0.50	0.42	0.49	0.10***
Middle father education (%)	0.29	0.46	0.33	0.47	−0.04***
High father education (%)	0.19	0.39	0.25	0.43	−0.06***
Stratum 1 (%)	0.36	0.48	0.25	0.43	0.12***
Stratum 2 (%)	0.36	0.48	0.37	0.48	−0.01**
Stratum 3 (%)	0.20	0.40	0.27	0.44	−0.07***
Stratum 4 (%)	0.05	0.22	0.07	0.26	−0.03***
Stratum 5 (%)	0.02	0.13	0.03	0.16	−0.01***
Stratum 6 (%)	0.01	0.10	0.02	0.14	−0.01***
Calendar A (%)	0.97	0.17	0.98	0.14	−0.01***
Students	4,673,332		12,647		4,685,979

Note: The natives are the students who reported that they were born in Colombia and the immigrants are those who were born in Venezuela. Math, Reading, Natural Sciences, Social Sciences and English correspond to the standardized scores obtained in the Saber 11 tests. Summary is the average standardized score of the five tests. The Stratum i variables correspond to the proportion of students living in a home in stratum i. The low mother-father education refers to incomplete secondary school; the middle category includes complete secondary and incomplete superior; and high category includes full superior or more. Calendar A correspond to students attending schools in calendar A, with school cycle between February and November.
Source: Own elaboration using Saber 11 tests from 2013 to 2022.

statistically significant.

In summary, Saber 11 data indicate that Venezuelan students generally underperform academically compared to their Colombian counterparts, despite their parents possessing higher educational levels. Predominantly, Venezuelan students are situated in urban areas. Additionally, it is evident that Venezuelan families typically occupy higher socioeconomic strata. According to data from the PERCE, prior to Hugo Chávez's presidency, Venezuela's education system achieved levels of education comparable to or lower than those of Colombia. This implies that the most educated Venezuelans, particularly those who have migrated to Colombia, are likely parents of students who exhibit higher socio-economic profiles compared to those of Colombian parents.

4 Identification strategy

In this paper, I exploit variations in the number of Venezuelan immigrants within schools, using the reopening of the border between Colombia and Venezuela in August 2016 as a natural experiment (Dustmann et al., 2016). Utilizing individual student score data, I analyze

school differences in the number of immigrants with Colombian students with a differences-in-differences (DID) approach. The treatment and comparison groups are determined based on whether schools received at least one Venezuelan student that is actually taking the Saber 11 tests following the border opening. A school is identified as "treated" if it has at least one Venezuelan student between 2017 and 2022, while "comparison" schools are those that did not receive any Venezuelan student during these years.

It is important to acknowledge that I cannot confirm whether school staff randomly allocate non-native students to different classes, or if class composition is influenced by the migration background of the students or other unobservable characteristics (Ammermueller & Pischke, 2009). To adjust for the potential non-random allocation of immigrants, I follow Schneeweis (2015) and Tonello (2016). This approach assumes that the relevant peer group is defined at the grade level rather than the class level. Consequently, my analysis concentrates on the variations in the 11th-grade composition of different cohorts within schools.

To estimate the average impact of this migration shock on the educational outcomes of native students, I estimate a standard DID model following Equation 1.

$$Y_{isdt} = \alpha + \beta Post_t * Treated_s + \delta X_{isdt} + \gamma_t + \gamma_d + \gamma_s + \epsilon_{isdt} \quad (1)$$

The variable of interest Y_{isdt} represents the standardized score achieved by native student i, at school s, located in the department d, in the year t. I estimate separate regressions for each competency, including Math, Reading, Natural Sciences, Social Sciences, and English, as well as a composite score of these five competencies. $Post_t$ is a dichotomous variable that takes the value 1 for years post-2016 and 0 for prior years. $Treated_s$ takes a value of 1 for treated schools and 0 otherwise. Additionally, I include year (γ_t), department (γ_d) and school (γ_s) fixed effects.[16] X_{isdt} is a set of control variables at both the student and school level, consisting of student gender and age, age squared, rural/urban residence, parent's education level, economic stratum, school size (measured by the number of students), the proportion of female students in the school, the academic calendar followed and the type of school day. Standard errors are clustered by school to correct for correlations in learning outcomes within the same school.

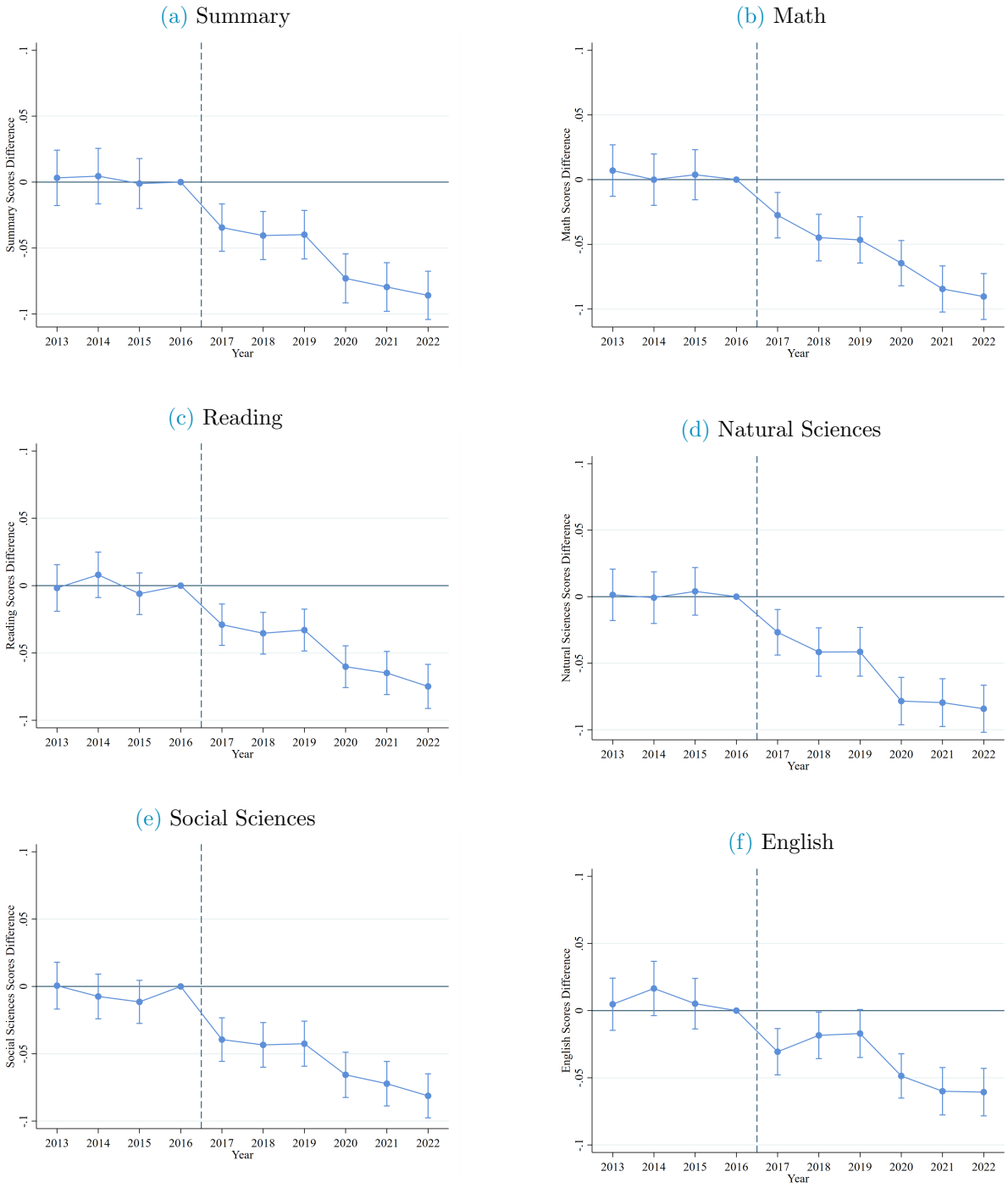
The coefficient of interest β is the one associated with the interaction between $Post_t$ and $Treated_s$, which identifies the peer effect of Venezuelan immigrants in the treated schools on the learning outcomes of Colombian students post-border opening, based on the parallel trends assumption. The parallel trends assumption requires that, in the absence of the treatment (migration of Venezuelans to Colombia), the evolution of the scores for treated natives would have been the same as the evolution for untreated natives. Although this assumption relies on a counterfactual and cannot be directly verified, I provide suggestive evidence to support it. Figure 4 illustrates the differences in scores between groups relative to the gap in 2016, accompanied by confidence intervals.

Prior to the border reopening, Figure 4 suggests that the trends between the treated and comparison groups were comparable. It is clear that for all of the outcomes cases, Math, Reading, Natural and Social Sciences, English, and Summary scores, the trends show no statistically significant differences.

Figure A.3 in the Appendix A shows the average scores for the six evaluated competencies, differentiated by treated schools (represented with a solid blue line) and comparison schools (represented with a dashed gray line). Initially, students in both groups exhibited similar performance levels. Following the border's reopening, however, schools in the treated group have generally underperformed relative to those in the comparison group, marking a significant divergence in trends.

[16] Note that school fixed effects account for non-random allocation of immigrants across classes, assuming that staffing decisions are consistent annually within each school.

Figure 4: Differences in average scores between treated and comparison students



Note: 95% confidence intervals are presented. Robust standard errors are clustered at the school level. Each panel of the figure presents the differences of the average standardized score between both, treatment and comparison groups. Treatment group are the students in the treated schools and the comparison group are the students in the remaining schools. The dark blue vertical line, with dashes, corresponds to the opening of borders between Colombia and Venezuela in mid-2016.
Source: Own elaboration using Saber 11 tests from 2013 to 2022.

Finally, to include additional context, Figure A.4 in the Appendix A shows the proportion of Venezuelan immigrants for the treated schools by years between 2013 and 2022. The Figure shows that before the opening of borders, between 2013 and 2016, there were few schools with some level of immigrants in the school, and this share greatly increased after 2017. Before

the opening of borders the highest observed level was around 20% of Venezuelans inside few schools and the period average between 2013 and 2016 was around 0.03%. After 2017 this number increased year by year with an average of 1.55% between 2017 and 2022, with a highest average value in 2022 of 3.56%.

5 Main results

5.1 Effect of immigration on learning of natives

Table 2 presents the main results from the estimation of Equation 1. Each row indicates the impact of Venezuelan immigration on the different learning outcomes of Colombian students. The coefficients reported correspond to β from Equation 1 for the different outcomes and specifications. Column 1 incorporates only the year, department and school fixed effects. Column 2 adds individual-level controls, including student gender and age, residential area, parental education levels and economic stratum. Column 3 includes school-level controls, such as total number of students, the proportion of female students and the type of school day.[17]

Table 2: Average effects of immigration on learning of native students

Score	(1)	(2)	(3)
Summary (β)	-0.014***	-0.016***	-0.018***
	(0.004)	(0.004)	(0.004)
Math (β)	-0.022***	-0.023***	-0.024***
	(0.004)	(0.004)	(0.004)
Reading (β)	-0.014***	-0.015***	-0.017***
	(0.004)	(0.004)	(0.004)
Natural Sci (β)	-0.019***	-0.020***	-0.022***
	(0.004)	(0.004)	(0.004)
Social Sci (β)	-0.015***	-0.017***	-0.019***
	(0.004)	(0.004)	(0.004)
English (β)	-0.005	-0.007	-0.008*
	(0.005)	(0.005)	(0.005)
Students	4,541,196	4,541,196	4,541,196
Schools	19,498	19,498	19,498
Year, Dept & School FE	Yes	Yes	Yes
Student controls	No	Yes	Yes
School controls	No	No	Yes

Note: *** ; ** ; * significant at the 1%, 5%, and 10% level, respectively. Robust standard errors by school clusters in parentheses. The natives are the students who reported that they were born in Colombia. Summary is the average standardized score between Math, Reading, Natural Sciences, Social Sciences and English. The reported coefficients are those that correspond to β from equation 1. Estimates in column 1 include only year, department and school fixed effects. Column 2 includes the fixed effects and individual controls (gender and age of student, area of residence, parents' educational level and economic strata). Column 3 includes fixed effects, individual controls, and school controls (total students, % of females in schools and school day).
Source: Own elaboration using Saber 11 tests from 2013 to 2022.

The results show a significant and negative effect across all evaluated competencies. The estimates suggest that students in schools that received at least one Venezuelan immigrant post-border opening scored, on average, 0.018 standard deviations lower than their peers in schools

[17] Table A.4 in the Appendix A provides the estimated coefficients for all variables included in the model.

without such immigrants. This effect size is smaller in English, with an effect of 0.008 standard deviations and bigger in Math with 0.024 standard deviations. Moreover, the magnitude of this effect appears to be negative in magnitude with the inclusion of individual- or school-level controls.

To contextualize and describe the economic significance of this finding, it is noted that students typically learn between 0.15 and 0.21 standard deviations over a standard school year (Evans & Yuan, 2019). Therefore, the effect I observe represents between 12% and 8.6% of a student's annual learning. Another way to contextualize this effect is to compare it to the relationship between maternal education and student achievement. The β coefficient of 0.018 reported in column 3 of table 2 corresponds to about 24.3% of the difference in standardized scores between native students whose mothers have low education (up to an incomplete secondary education) and whose whose mothers have middle education (between a complete secondary education and an incomplete superior education), where the coefficient is 0.074. In comparison, this effect size amounts to roughly 8.3% of the differential observed between students with low-educated and high-educated mothers (coefficient of 0.218).

5.2 Robustness checks

In this section, I conduct robustness checks to enhance the validity of my findings. First, I utilize a balanced panel of schools, those consistently observed throughout the analysis period (between 2013 and 2022). This approach aims to determine whether the observed effects are influenced by schools opening or closing during the study period. 54.3% of the students are enrolled in schools that are consistently present over the ten years.[18] The results with the balanced panel estimation of schools show similar coefficients (refer to column 2 of Table 3).

For the second robustness check, I excluded students from Bogotá and Antioquia, the two departments with the largest populations in Colombia. Bogotá consistently achieved the highest scores across all tests, while Antioquia's scores were above average. The results show smaller effects (column 3), and for some cases they are statistically different compared to the baseline estimates (column 1).

The third check, is to include students from departments not covered by the GEIH as additional observations (column 4). Of these, 60.4% belong to the comparison group and 39.6% to the treatment group. This adjustment increases the sample size by approximately 131,000 observations (2.8%), comprising students included as a comparison and treatment group, based on the immigration status of their schools post-border opening.[19] The estimates from this expanded dataset are similar to baseline results.

The fourth check included schools following Calendar B, which includes 2.9% of the student population. These schools operate from September to June, differing in testing schedule from Calendar A schools. Schools in Calendar B also report higher average scores, around 0.97 standard deviations, whereas schools in Calendar A have an average of 0.01 standard deviations. The inclusion of Calendar B schools yields estimates slightly upward the baseline, however, these differences are not statistically significant.

The fifth check involved omitting test scores from 2020. The ICFES administered the 2020 tests via phone calls during the COVID-19 pandemic, significantly reducing the test partici-

[18]The distribution of students across schools with varying operational durations is as follows: 4.5% in schools operating for 9 years, 3% for 8 years, 10.5% for 7 years, 9.5% for 6 years, 5% for 5 years, 6.5% for 4 years, 5.2% for 3 years, and 1.4% for schools operating for just 2 and 1 years combined.

[19]Students from these departments (Amazonas, Arauca, Casanare, Guainia, Guaviare, Putumayo, Vaupés and Vichada) performed below the national average.

Table 3: Robustness checks: Sample modifications

Score	Base	Schools	Dept Bog-Ant	Dept Saber	Calendar	Year 2020	School Pre
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Summary (β)	-0.018***	-0.015***	-0.010**	-0.017***	-0.022***	-0.017***	-0.017***
	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Math (β)	-0.024***	-0.022***	-0.021***	-0.024***	-0.027***	-0.025***	-0.023***
	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Reading (β)	-0.017***	-0.016***	-0.009**	-0.016***	-0.019***	-0.016***	-0.016***
	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
Natural Sci (β)	-0.022***	-0.023***	-0.012**	-0.021***	-0.024***	-0.020***	-0.021***
	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Social Sci (β)	-0.019***	-0.023***	-0.008*	-0.019***	-0.021***	-0.019***	-0.020***
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
English (β)	-0.008*	-0.008	-0.012**	-0.008*	-0.015***	-0.010**	-0.007
	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
Students	4,541,196	2,439,583	3,163,581	4,671,441	4,673,166	4,126,437	4,443,349
Schools	19,498	6,059	14,835	20,242	19,936	19,366	19,319
FE & Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: *** ; ** ; * significant at the 1%, 5%, and 10% level, respectively. Robust standard errors by school clusters in parentheses. The natives are the students who reported that they were born in Colombia. Summary is the average standardized score between Math, Reading, Natural Sciences, Social Sciences and English. The reported coefficients are those that correspond to β from equation 1. All estimates include the same controls and fixed effects as column 3 of Table 2. Column 1 shows the results obtained in column 3 of Table 2. Column 2 shows the results with a full panel sample of schools. Column 3 shows the results excluding the two departments with the largest population according to the GEIH, Bogotá and Antioquia. Column 4 shows the estimate increasing the number of students by including the departments not surveyed by the GEIH (Amazonas, Arauca, Casanare, Guainia, Guaviare, Putumayo, Vaupés and Vichada). Column 5 shows the estimate including schools in calendar B and column 6 excludes year 2020. Column 7 shows results excluding schools with immigrants before the borders opening.
Source: Own elaboration using Saber 11 tests from 2013 to 2022.

pation rate.[20] This testing method could introduce some noise to my estimations. Estimates with this modification of sample are similar to baseline results.

The sixth and final robustness check consists in excluding schools that received Venezuelan immigrants before the opening of borders. These schools corresponds, in total, to less than 1% of schools between 2013 and 2016 and are excluded from all the periods of analysis. Column 7 shows estimates statistically similar results than baseline results presented in column 1.

Table 3 also shows the same robustness checks for the five individual scores: Math, Reading, Natural Sciences, Social Sciences, and English. The results indicate that the coefficients are generally consistent with the baseline estimates presented in column 1, with a couple of notable exceptions. Firstly, the coefficients for Reading and both Natural and Social Sciences during the robustness check involving the exclusion of larger departments (column 3) are significantly lower than those in the baseline. This variance suggests that the effects observed for Reading, and Natural and Social Sciences might be particularly sensitive to the exclusion of schools from heavily populated departments, which typically have higher average scores or larger school sizes.

A second exception is the English coefficient for the robustness checks in column 5, the coefficient is not only bigger, but it also is slightly significant. Given that some Calendar

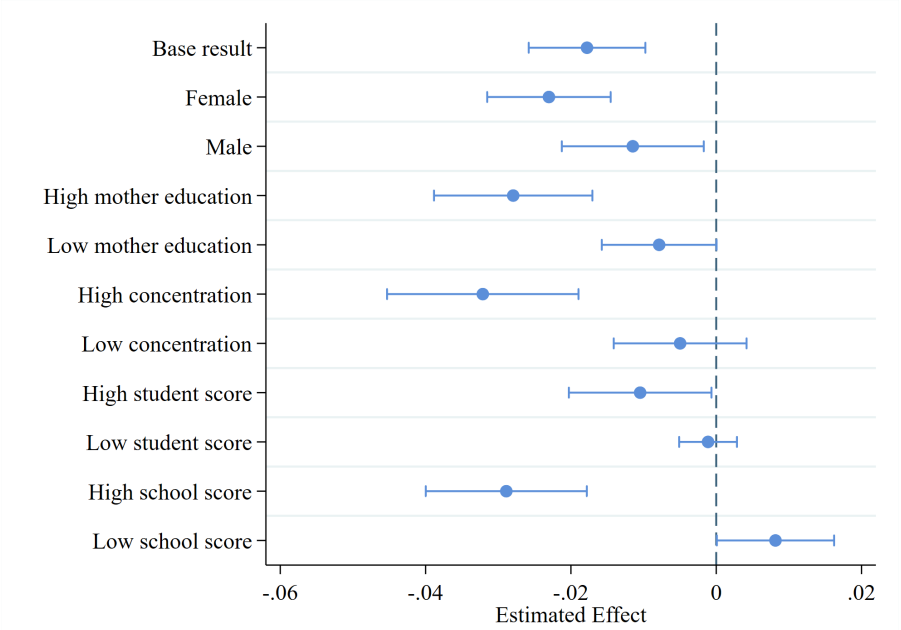
[20]From 2013 and 2019, an average of 485,600 students took the test annually. In 2020, this number decreased by approximately 13% to 424,050.

B schools offer classes in English and follow the North American and European academic calendars, it is likely that the results are sensitive to the inclusion of these schools, which generally have higher levels of English proficiency.[21]

5.3 Heterogeneous effects

Up to this point, I have assumed that the effect of the presence of Venezuelans is uniform for all native students. However, existing literature indicates that this effect might vary among different groups of natives based on their socioeconomic status, gender or prior academic performance. To determine if this is the case in this context, I examine the effect of immigrant presence for different groups of natives by running separate estimates for each group. I estimate heterogeneous effects for the following categories: gender (men and women); parental education level (natives with mothers of varying educational backgrounds); school environment (natives in schools with differing concentrations of highly educated mothers); academic standing (natives with high versus low performance); and school performance level (natives in schools with varying levels of average academic achievement).

Figure 5: Heterogeneous effects by different groups of native students



Note: 95% confidence intervals are presented. Robust standard errors are clustered at the school level. The reported coefficients are those that correspond to β from equation 1 for each of the groups. The effects of *High mother education* and *Low mother education* are estimates for natives who have mothers with high and medium educational levels and for natives with mothers with low education, respectively. The effects of *High concentration* are for natives who attend schools with more than 50% of mothers with high and middle education and *Low concentration* are for natives in schools with less than 50% of mothers with high and middle education. The *High student score* and *Low student score* estimates are for native students who obtain scores above and below the median year by year, respectively. The coefficients for *High school score* and *Low school score* are for native students attending schools with average scores above and below the school median year-to-year, respectively.
Source: Own elaboration using Saber 11 tests from 2013 to 2022.

Figure 5 presents the estimated values from Equation 1 focusing on the Summary score across the specified groups. The results indicate some differences between male and female

[21]Schools in calendar A have an average English score of -0.005 standard deviations, whereas schools in calendar B average 1.24 standard deviations.

students, with women experiencing a slightly higher negative effect compared to men. Additionally, the negative impact is more pronounced among students whose mothers have higher educational levels and those attending schools where a high proportion of mothers are highly educated (defined as schools where over 50% of mothers have advanced education). Similarly, students with above-median personal scores and those in higher-performing schools (schools whose average score is above the median) also exhibit a greater negative effect. Conversely, the effects are not significant for students in schools with a low proportion of highly educated mothers and for those with below-median scores. The effect is positive, but slightly, significant for students in low-performing schools.[22]

The summary score results are generally similar to those found for the five scores presented individually in the Figure A.5 in the Appendix A. Male and Female difference is significant for all scores except Reading and English. The effect is more negative for students whose mothers have higher educational backgrounds and less so for those from backgrounds with low education. The negative effect turns not-significant for students attending schools with low proportion of highly educated mothers for all outcomes except Math. Furthermore, performance-based disparities are consistently significant, with low-performing students and those from lower-performing schools experiencing less pronounced negative impacts. However, these effects are not significant.[23]

These results highlight significant heterogeneity based on the gender, the academic scores of natives, the school environment, and their mothers' educational levels. In general, Colombian natives experience a disadvantage from the presence of Venezuelan immigrants. Notably, this effect is less negative for natives from lower socio-economic backgrounds. Conversely, the negative effect is more pronounced for students with highly educated mothers, those attending schools with a high proportion of educated mothers, and schools with above-average academic performance. These heterogeneous effects indicate that Venezuelan immigrants disproportionately affect high-achieving students, students in high-achieving schools, and those whose mothers have higher educational levels.

6 Mechanisms and extensions

A natural first question is about the mechanisms that likely generate these negative effects. While a comprehensive exploration of all possible channels is beyond the scope of this paper, I am able to explore a channel. Until now, I have assumed that the effect of the presence of Venezuelans is independent of the total number or proportion of immigrants in the school. As is common in this literature, papers typically use variables that measure the proportion of immigrants at various levels of aggregation, for example, by region[24], by school[25], by grade[26] or by class[27]. To observe the intensive effects of a 1% increase in immigrant students with a

[22] I also estimated heterogeneities according to the area of residence and found that negative effect for urban schools is higher, but the native and immigrant population in rural area is low, so I decided to not include that result.

[23] The case for English is distinct, the effect is positive and significant for students with mothers with low education and low performance, and students attending schools with low proportion of educated mothers and with low performance. Additionally, the negative impacts for high-performing students and men become non-significant.

[24] For example, Hunt (2017); Luksic (2022); van der Werf (2021)

[25] For example, Hermansen and Birkelund (2015); Jensen and Rasmussen (2011); Pedraja-Chaparro et al. (2016)

[26] For example, Bossavie (2020); Fletcher et al. (2021); Gould et al. (2009); Schneeweis (2015); Tonello (2016)

[27] For example, Ballatore et al. (2018); Contini (2013); Figlio et al. (2021); Frattini and Meschi (2019); Ohinata and Van Ours (2013)

DID design, I follow a procedure similar to Pedraja-Chaparro et al. (2016). Based on equation 1, I include an additional interaction between the DID design variable and the proportion of Venezuelan immigrants per school (Schneeweis, 2015; Tonello, 2016) and the proportion by itself, thereby I estimate Equation 2.

$$\begin{aligned} Y_{isdt} =& \alpha + \beta Post_t * Treated_s + \mu \%Migrants + \lambda Post_t * Treated_s * \%Migrants \\ &+ \delta X_{isdt} + \gamma_t + \gamma_d + \gamma_s + \epsilon_{isdt} \end{aligned} \tag{2}$$

Y_{isdt}, X_{isdt}, γ_t, γ_d, γ_s and ϵ_{isdt} are defined similar as Equation 1. β coefficient is still the one associated with the DID design, identifying the peer effect in the treated schools after the border opening. μ coefficient quantifies the effect of a 1% increase in the proportion of immigrants within a school. And λ coefficient is associated with the peer effect in the treated schools after the opening of borders if the school were to increase the proportion of Venezuelan immigrants by 1%.

A second step is to find evidence if these intensive effects are nonlinear to understand if this negative effect could be even higher when the immigrant share at grade change. To this purpose, I introduce a quadratic term for the immigrant share at the grade level This approach aims to discern nonlinear dynamics in how immigrant shares influence native students' learning outcomes.Thus, I estimate equation 3.

$$\begin{aligned} Y_{isdt} =& \alpha + \beta Post_t * Treated_s + \mu_1 \%Migrants + \mu_2 \%Migrants^2 \\ &+ \lambda_1 Post_t * Treated_s * \%Migrants + \lambda_2 Post_t * Treated_s * \%Migrants^2 \\ &+ \delta X_{isdt} + \gamma_t + \gamma_d + \gamma_s + \epsilon_{isdt} \end{aligned} \tag{3}$$

Results for estimates of equation 2 and equation 3 include the same controls and fixed effects as column 3 of Table 2. Panel A of Table 4 present the coefficients for the linear effects (equation 2) and Panel B present the coefficients for the nonlinear effects (equation 3).

The estimates confirm that the peer effect, measured by β, is similar as the estimates from column 3 of Table 2, showing a reduction of 1.8% of a standard deviation. The intensive effect in Panel A, measured by λ, indicates that native students in treated schools experiencing a 1% increase in Venezuelan immigrants post-border opening see an increase in scores by 0.012 standard deviations. This intensive effect is similar in magnitude for Math and for the remaining outcomes the effects are not significant. Panel B shows nonlinearity in the effects based on the share of immigrants, linear share coefficient (λ_1) is positive, while the squared share coefficient (λ_2) is negative, and both of them are non-significant for all of the outcomes.

These findings suggest that the presence of Venezuelan immigrants has a detrimental effect on the learning outcomes of native students, as compared to the comparison group. The introduction of the immigrant share indicates that the negative effect can turn to zero if the proportion of immigrants increases.

It can be argued, that after the inclusion of proportion of immigrants within the school, native students are negatively affected from having an increased presence of immigrants with them. Based on the statistics presented in Table 1, I can believe that teachers need to allocate more class time to assist Venezuelan immigrant students catch up to their peers, as they tend to have lower academic performance. This negative effect represents an indirect peer effect, where the spillover is predominantly influenced by teacher attitudes.[28]

[28]This mechanism with negative effect is similar to the findings of Bossavie (2020).

Table 4: Intensive effects of immigration on learning of native students

Variables	Summary (1)	Math (2)	Reading (3)	Natural Sci (4)	Social Sci (5)	English (6)
Panel A: Equation 2						
School (β)	-0.018***	-0.024***	-0.016***	-0.022***	-0.019***	-0.008*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
%immigrants (μ)	-0.011*	-0.013**	-0.007	-0.004	-0.003	-0.012
	(0.006)	(0.006)	(0.005)	(0.006)	(0.005)	(0.008)
School $*$ *%immigrants* (λ)	0.012*	0.013**	0.006	0.004	0.004	0.012
	(0.006)	(0.006)	(0.005)	(0.006)	(0.005)	(0.008)
Panel B: Equation 3						
School (β)	-0.018***	-0.024***	-0.016***	-0.021***	-0.019***	-0.008*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
%immigrants (μ_1)	-0.014	-0.016*	-0.006	-0.005	-0.007	-0.017
	(0.009)	(0.010)	(0.007)	(0.009)	(0.007)	(0.011)
%immigrants^2 (μ_2)	0.001	0.001	-0.000	0.000	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
School $*$ *%immigrants* (λ_1)	0.015	0.016	0.005	0.004	0.007	0.018
	(0.009)	(0.010)	(0.007)	(0.009)	(0.007)	(0.011)
School $*$ *%immigrants*^{2} (λ_2)	-0.001	-0.001	0.000	-0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Students	4,541,196	4,541,196	4,541,196	4,541,196	4,541,196	4,541,196
Schools	19,498	19,498	19,498	19,498	19,498	19,498
FE & Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: *** ; ** ; * significant at the 1%, 5%, and 10% level, respectively. Robust standard errors by school clusters in parentheses. The natives are the students who reported that they were born in Colombia. Summary is the average standardized score between Math, Reading, Natural Sciences, Social Sciences and English. The reported coefficients in Panel A are those that correspond to β, λ and μ from equation 2. The reported coefficients in Panel B are those that correspond to β, λ_1, λ_2, μ_1 and μ_2 from equation 3. All estimates include the same controls and fixed effects as column 3 of Table 2.
Source: Own elaboration using Saber 11 tests from 2013 to 2022.

Conversely, if the immigrant share within grades is sufficiently large, these negative effect may disappear, or, the negative effect is even higher at low levels of immigrant concentration.[29] One possible explanation is that, if the immigrant share within grades becomes sufficiently large, it may become economically viable for schools to hire additional teachers or establish more classes specifically for immigrants. This adaptation could mitigate the disproportionate allocation of class time, thereby diminishing the negative spillovers experienced by native students. Another possible explanation is that schools redirect resources due to an increase of low-performing immigrants (e.g., reallocation of better teachers or segregation of immigrants in different classes).

A second question concerns the variation in the impact as native students' exposure to immigrants in schools increases over time. To answer this, I employ an event study design that captures trends from up to eight years before and six years after the first Venezuelan immigrant arrives in 11th grade, following the mid-2016 border opening. Due to few observations between eight and five years before the arrival of immigrants, I group them into a single group. This approach not only provides insights into the dynamics of the immigrant effect on native scores,

[29]Similar result to the one found by Green and Vaag Iversen (2022) and by Gould et al. (2009) on matriculation instead of learning.

but also allows me to further validate the DID results. To estimate this variation in the effect of the migration shock on native scores, I estimate the specification of Equation 4.

$$Y_{isdt} = \alpha + \sum_{j=2}^{5} \beta_{-j} * D_{s,-j} + \sum_{j=0}^{5} \beta_j * D_{s,j} + \delta X_{isdt} + \gamma_t + \gamma_d + \gamma_s + \epsilon_{isdt} \quad (4)$$

Y_{isdt}, X_{isdt}, γ_t, γ_d, γ_s and ϵ_{isdt} are defined similar as Equation 1. $D_{s,-j}$ are dummy variables that identify one of the four years before the previous year of immigrant arrival to school s and $D_{s,j}$ identifies one of the six years after the arrival of immigrants, with $j = -1$ (the year just before the arrival of the first immigrant) serving as the omitted category. The coefficients β_j quantify the dynamic effects of immigrant presence in treated schools for each year j post-arrival, relative to the baseline period. In contrast, β_{-j} coefficients capture previous trends prior the event. β coefficients are estimated with a comparison group of native students attending schools without immigrant presence, the same comparison group as in the DID approach.

Figure 6 presents the results for the summary score as specified in Equation 4. This figure delineates the variation in effects over time on native scores due to the presence of immigrants, highlighting how these effects evolve with increased exposure to immigrant students. The vertical dashed line indicates the moment of treatment, the arrival of immigrants, while the horizontal dashed line corresponds to the baseline effect estimated in Table 2 for the summary score. The red line is the average coefficient before treatment with the 95% confidence interval in short dashed lines. This helps me to evaluate a conjoint hypothesis of pre-trends for all of the coefficients at the same time. Table A.5 in the Appendix A presents the estimates of each coefficient for the six outcomes, including the estimate and significance of the main results, and the average of the five coefficients pre-treatment and significance to test for parallel trends.

The results confirm the negative impact on native students' scores of Table 2, with the magnitude of this effect remaining more or less constant around the effect of the main results. This negative effect is both persistent and significant from the first three to five years natives share schools with immigrants, exhibiting an increase of confidence intervals between years four to six turning the effect non-significant. Column 1 of Table A.5 in the Appendix A presents each of the coefficients for the summary outcome and I can see that between the first three years the negative effect is, on average, -0.012 standard deviations, increasing to -0.030 in the fifth year, and turning non-significant by the sixth year. This results show that negative effect tends to disappear as native students exposition to immigrants increases. This suggests that schools, teachers and native students adapt to immigrants in schools and better adapt coexistence in schools, classrooms and pedagogical practices.

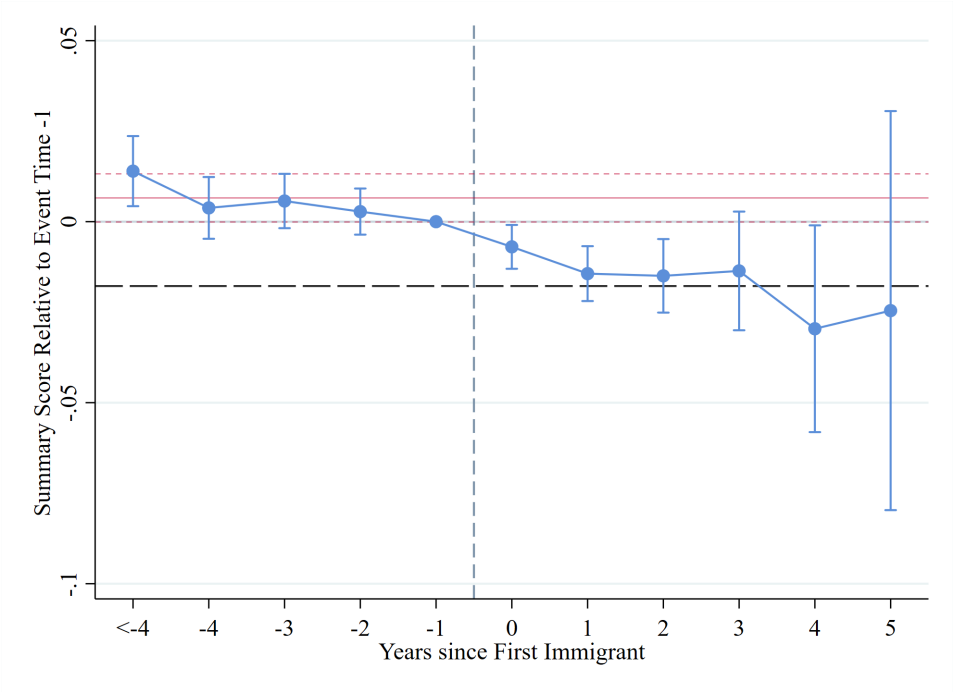
The summary score results are similar to those observed for the math and natural sciences scores, as presented individually in Figure A.6 in the Appendix A. I observe negative effects for the first three to four years, and, after that, effects tend to zero. The only exception is Math, where the effect for the fifth year is -0.073 standard deviations.[30]

Finally, a third question is about exposure of native students to Venezuelan immigrants not only inside the school. Figure A.7 in the Appendix A illustrates the distribution of Venezuelan immigrants across Colombian departments 12 months after the border reopening. The Figure highlights regions with the highest and lowest influx of Venezuelan migrants.[31] To capture

[30]Note that in this case I can only reject non differential trends between treatment and comparison groups at 90% level. The only exception is Natural Sciences where I have different trends between groups before treatment.

[31]Figure A.7 also presents the proportion of Venezuelans aged 12-17 and 16-17, by department, relative to the total population of these age groups, 12 months post-border opening. It reveals that the migratory intensity for adolescents and those at the typical age for taking the Saber 11 test (16 to 17 years old) is similar to the overall migration patterns to these departments.

Figure 6: Event study of the effects of immigration on learning of native students



Note: 95% confidence intervals are presented. Robust standard errors are clustered at the school level. The reported coefficients are those that correspond to β_{-j} and β_j from equation 4. The dark blue vertical line, with dashes, corresponds to the moment of treatment of schools. The dark blue horizontal line, with dashes, shows the results obtained in column 3 of Table 2 for the summary score. The red line represents the average of pre-treatment coefficient and the red dashed lines are the confidence intervals for that average.
Source: Own elaboration using Saber 11 tests from 2013 to 2022.

the additional exposure of natives, I create a second treated group based on the departmental proportion of Venezuelan immigrants. Utilizing the 95th percentil of departments by Venezuelan population as a threshold, I categorize departments above this threshold, such as *Norte de Santander* and *La Guajira*, as the treated group due to their higher immigrant influx, while the remaining 22 departments serve as comparisons. To incorporate this additional treatment to the analysis, I employ a triple differences (DDD) design as specified in Equation 5.

$$\begin{aligned} Y_{isdt} = \alpha + \beta Post_t * Treated_s + \mu Post_t * Dept_d + \lambda Post_t * Treated_s * Dept_d \\ + \delta X_{isdt} + \gamma_t + \gamma_d + \gamma_s + \epsilon_{isdt} \end{aligned} \tag{5}$$

Y_{isdt}, X_{isdt}, γ_t, γ_d, γ_s and ϵ_{isdt} are defined similar as Equation 1. $Dept_d$ is introduced to differentiate between treated departments, with *Norte de Santander* and *La Guajira* assigned a value of 1, and all other departments a value of 0. β is still the associated coefficient with the DID design from the main results, identifying the effect in treated schools after the opening of borders. μ is the coefficient that measures the effect of Venezuelan immigration on native students' scores across these treated departments post-border opening. Additionally, λ identifies the peer effect of Venezuelan immigrants in schools that are not only located in treated departments, but have also directly received immigrants after the border was opened.

Table 5 presents estimates from Equation 5 for the summary score and the five individual scores, alongside the sum of β or μ with λ in brackets. Firstly, the results for treated schools post-border opening (β) are negative and significant, similar to the results from column 3 of Table 2. Secondly, a positive and significant effect is observed for students in departments with

Table 5: Triple differences effects of immigration on learning of native students

Variables	Summary (1)	Math (2)	Reading (3)	Natural Sci (4)	Social Sci (5)	English (6)
School (β)	-0.018***	-0.024***	-0.016***	-0.022***	-0.019***	-0.006
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
	[-0.040]**	[-0.034]	[-0.036]**	[-0.025]	[-0.025]	[-0.060]***
Department (μ)	0.073***	0.063***	0.067***	0.064***	0.067***	0.047***
	(0.011)	(0.013)	(0.011)	(0.012)	(0.011)	(0.014)
	[0.050]***	[0.053]***	[0.048]***	[0.061]***	[0.061]***	[-0.007]
School $*$ *Department* (λ)	-0.022	-0.010	-0.019	-0.003	-0.006	-0.054**
	(0.018)	(0.023)	(0.016)	(0.018)	(0.017)	(0.022)
Students	4,541,196	4,541,196	4,541,196	4,541,196	4,541,196	4,541,196
Schools	19,498	19,498	19,498	19,498	19,498	19,498
FE & Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: *** ; ** ; * significant at the 1%, 5%, and 10% level, respectively. Robust standard errors by school clusters in parentheses. The addition between β or μ and λ coefficients in brackets. The natives are the students who reported that they were born in Colombia. Summary is the average standardized score between Math, Reading, Natural Sciences, Social Sciences and English. All estimates include the same controls and fixed effects as column 3 of Table 2.
Source: Own elaboration using Saber 11 tests from 2013 to 2022.

a high immigrant share, post-border opening, where the average effect is 7.3% of a standard deviation. This effect is consistently positive across all individual scores. Finally, there is no significant effect observed for treated schools withing treated departments when compared to untreated schools in untreated departments, with the exception of the English score, where the effect is -0.05 standard deviations.

Moreover, when considering specifically students residing in treated departments, the positive effect diminishes for those attending treated schools. For natives in *Norte de Santander* and *La Guajira*, the positive effects reduces from 0.07 to 0.05 standard deviations for those attending treated schools. This result persists across all individual scores, except for English, where the effect is precisely zero. These results indicate that the negative peer effect I identified in the previous section, becomes positive when I include migration patterns into the analysis, particularly for students residing in departments with a high proportion of Venezuelan immigrants.

From these findings I can believe three things. First, as treated departments tend to have lower average scores, differences between native and immigrant students are not as large as differences from Table 1. To verify this, I present Table A.6 in the Appendix A, which indicates that there are no significant performance differences between native and immigrant students in the treated departments across summary, natural sciences, and English scores. However, there are negative differences in reading and social sciences, and positive differences in math. Conversely, in comparison departments, there are positive and significant differences in favor of natives. This suggests that the negative indirect peer effects I found, presumably driven by teacher attitudes and class time allocation, are mitigated when performance levels between native and immigrant groups are comparable.

Second, it is difficult to verify this claim with the available data and it will require further research. However, I can believe that native students in regions with a higher concentration of immigrants might be better adapted to such demographics compared to their counterparts in areas with fewer immigrants. Consequently, when native students interact with immigrants in

schools, their familiarity with such interactions could foster mutual benefits, particularly when there are negligible differences in academic performance between the groups.

Third, this results could suggest that Colombian departments applied differentiated policies for immigrants between the years of analysis. Between those policies there could be pro-immigrant policies and policies to prepare schools for immigrants arrival. I will discuss this more in the next section.

7 Limitations and next steps

In this section I present some limitations of the study and propose next steps to continue working on the agenda. Regarding limitations, first, treatment definition is not as clean as in other literature because it relies on observing Venezuelans taking the Saber 11 test and not attending schools. Based on Figure 1, I can conclude that not all Venezuelans living in Colombia are actually taking the test or are enrolled in secondary education. The next step is to estimate enrollment rates of immigrants and natives, compare them to the number of students taking the test year by year, and try to understand if Venezuelans are attending schools and not taking the test, or if they are not enrolled in school.

Second, the treatment I propose is based on the observation of at least one Venezuelan student in schools between 2017 and 2022 without distinction of the year. This could bias the results as I am pooling schools treated at 2017 and schools treated at 2022 in the same treatment group. A subsequent step to address this is to define different treatments and different comparison groups. One approach is to define treatment on a yearly basis, assuming that once a school is treated, it remains treated, and then estimate differentiated effects by year between 2017 and 2022. I can also define only two treatment groups: an early treated group that received immigrants between 2017 and 2019, and a late treated group that received immigrants between 2020 and 2022. Another approach is to restrict the sample to only treated schools that differed in the time they were treated and make comparisons between an early treated group and a late treated group. This approach is intriguing because school selection bias might be lower between early treated and late treated schools compared to schools that received immigrants and those that did not.

Third, a possible mechanism of the effects could be parents' labor markets outcomes instead of school peer effects. High rates of immigration could negatively affect employment of native parents, causing a decrease in household income, and ultimately affecting learning. Another limitation related to this is regional trends, as Colombian municipalities or departments that are preferred by immigrants to establish due to unobservable characteristics. To control for these regional trends in my estimates, I plan to include high-dimension fixed effects combining municipality and year.

Fourth, the triple differences results presents some positive effects for natives learning of immigrants arrival to departments with the highest influx of Venezuelans. This result can suggest me that the population of those departments adapted better to the big influx of immigrants or that those departments implemented specific policies to better receive these immigrants. My next step is to identify pro-immigrant policies in Colombia between the years of my analysis and to compare the policies between departments.

8 Conclusions

This paper investigates the impact of the Venezuelan migratory exodus on the educational outcomes of Colombian natives high school students. Utilizing data from the Saber 11 tests

and the 2016 reopening of borders between Venezuela and Colombia, I analyze the impact using a difference-in-differences design. To identify comparison and treatment groups, I use an immigrant concentration within schools approach, defining as treated schools those that received at least one Venezuelan after the opening of borders.

The results reveals a consistent negative effect on the academic performance of native students within treated schools with a magnitude of 1.8% of a standard deviation decrease (or -0.018 standard deviations). This detrimental effect is evident from the first year following the immigrant's arrival. Notably, the decline in performance disappear between the fourth and fifth years, except for Math, where the negative effect gets bigger. The negative effects are somewhat robust to the use of different samples.

Heterogeneous effects suggests large variations between groups. Natives with mothers that are highly educated or students that had higher performance experience higher negative effects than their counterparts. It is also noted that natives attending schools with high average performance and with a high concentration of educated mothers, also experience higher negative effects than their counterparts. There is also a higher negative effect for women compared to men. For students attending schools with an average low score or for students with low performance the effect tends to disappear and be precisely zero.

Intensive effects suggests that an increase of 1% in the number of immigrants in treated schools increases natives learning an extra 0.01% of a standard deviation. As Venezuelan immigrants tend to have lower academic performance, I believe that teachers need to allocate more class time to help them, causing an indirect negative effect over native students. Despite this negative mechanism, the effect disappear as the share of immigrants increases. When the concentration of immigrants is lower, the negative effect is still significant. The negative effect also disappears when the academic performance of natives and immigrants is similar, and it becomes positive when the share of immigrants per department is high.

References

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.

Bahar, D., Ibáñez, A. M., & Rozo, S. V. (2021). Give me your tired and your poor: Impact of a large-scale amnesty program for undocumented refugees. *Journal of Development Economics*, *151*, 102652.

Ballatore, R. M., Fort, M., & Ichino, A. (2018). Tower of Babel in the classroom: Immigrants and natives in Italian schools. *Journal of Labor Economics*, *36*(4), 885–921.

Bossavie, L. (2020). The effect of immigration on natives' school performance does length of stay in the host country matter? *Journal of Human Resources*, *55*(2), 733–766.

Caruso, G., Canon, C. G., & Mueller, V. (2019). Spillover effects of the Venezuelan crisis: migration impacts in Colombia. *Oxford Economic Papers*, *73*(2), 771–795.

Contini, D. (2013). Immigrant background peer effects in Italian schools. *Social science research*, *42*(4), 1122–1142.

Contreras, D., & Gallardo, S. (2022). The effects of mass migration on the academic performance of native students. Evidence from Chile. *Economics of Education Review*, *91*, 102314.

DANE. (2021). *Población migrante venezolana en Colombia, un panorama con enfoque de género* (Tech. Rep.). Nota estadística, Departamento Administrativo Nacional de Estadística.

Delgado-Prieto, L. (2021). Dynamics of local wages and employment: Evidence from the Venezuelan immigration in Colombia. *Universidad Carlos III de Madrid, Departamento de Economía, working paper*, *31183*.

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.

Entorf, H., & Lauk, M. (2008). Peer effects, social multipliers and migrants at school: An international comparison. *Journal of Ethnic and Migration Studies*, *34*(4), 633–654.

Eslava, M., Revolo, M. J., & Ortiz, R. P. (2021). Diseño y formulación de subsidios a la demanda de energía eléctrica en Colombia (Design and formulation of subsidies for electricity demand in Colombia). *Documento CEDE*(15).

Evans, D., & Yuan, F. (2019). Equivalent years of schooling: A metric to communicate learning gains in concrete terms. *World Bank Policy Research Working Paper*(8752).

Figlio, D., Giuliano, P., Marchingiglio, R., Özek, U., & Sapienza, P. (2021). *Diversity in schools: Immigrants and the educational performance of US born students* (Tech. Rep.). National Bureau of Economic Research.

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.

Fletcher, J., Kim, J., Nobles, J., Ross, S., & Shaorshadze, I. (2021). The effects of foreign-born peers in us high schools and middle schools. *Journal of Human Capital*, *15*(3), 432–468.

Frattini, T., & Meschi, E. (2019). The effect of immigrant peers in vocational schools. *European Economic Review*, *113*, 1–22.

Geay, C., McNally, S., & Telhaj, S. (2013). Non-native speakers of English in the classroom: what are the effects on pupil performance? *The Economic Journal*, *123*(570), F281–F307.

Gould, E. D., Lavy, V., & Paserman, D. (2009). Does immigration affect the long-term educational outcomes of natives? Quasi-experimental evidence. *The Economic Journal*, *119*(540), 1243–1269.

Green, C., Haaland, K., & Vaag Iversen, J. (2022). Immigrant peers and foreign language acquisition. *IZA Discussion Paper*.

Green, C., & Vaag Iversen, J. (2022). Refugees and the educational attainment of natives: Evidence from norway. *Economics of Education Review*, *88*, 102258.

Hermansen, A. S., & Birkelund, G. E. (2015). The impact of immigrant classmates on educational outcomes. *Social Forces*, *94*(2), 615–646.

Hernandez, I., & Monaldi, F. (2016). Weathering collapse: An assessment of the financial and operational situation of the Venezuelan oil industry. *CID Working Paper Series*.

Hunt, J. (2017). The impact of immigration on the educational attainment of natives. *Journal of Human Resources*, *52*(4), 1060–1118.

Ibánez, A. M., Moya, A., Ortega, M. A., Chatruc, M. R., & Rozo, S. V. (2021). Life out of the shadows: Impacts of amnesties in migrant's life. In *2nd research conference on forced displacement.*

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 Review*, *30*(6), 1503–1515.

Lebow, J. (2021). The labor market effects of Venezuelan migration to Colombia: Reconciling conflicting results. *IZA Journal of Development and Migration*, *13*(1).

Lombardo, C., & Peñaloza-Pacheco, L. (2021). Exports "brother-boost": The trade-creation and skill-upgrading effect of Venezuelan forced migration on Colombian manufacturing firms. *Documentos de Trabajo del CEDLAS*.

Luksic, J. D. (2022). *Can immigration affect place effects? Accounting for the indirect effects*

of immigrants on native test scores. Retrieved from https://sites.google.com/view/juandiegoluksic/

Martínez, D., & Martínez, M. (2023). Immigrant children and optimal school choice: Evidence from the Venezuelan migration to Peru. *Job Market Paper*. Retrieved from https://acsweb.ucsd.edu/~djmartin/pdfs/draftJMP.pdf

Ministerio de Educación. (2015). *Decreto No. 1075* (Tech. Rep.).

Morales, C. (2022). Do refugee students affect the academic achievement of peers? Evidence from a large urban school district. *Economics of Education Review*, *89*, 102283.

Ohinata, A., & Van Ours, J. C. (2013). How immigrant children affect the academic achievement of native Dutch children. *The Economic Journal*, *123*(570), F308–F331.

Ohinata, A., & Van Ours, J. C. (2016). Quantile peer effects of immigrant children at primary schools. *Labour*, *30*(2), 135–157.

Pedraja-Chaparro, F., Santín, D., & Simancas, R. (2016). The impact of immigrant concentration in schools on grade retention in spain: a difference-in-differences approach. *Applied Economics*, *48*(21), 1978–1990.

Pedrazzi, J. P., & Peñaloza-Pacheco, L. J. (2020). Heterogeneous effects of forced migration on female labor supply. *Documentos de Trabajo del CEDLAS*.

Peñaloza-Pacheco, L. (2019). Living with the neighbors: The effect of Venezuelan forced migration on wages in Colombia. *Documentos de Trabajo del CEDLAS*.

Romero, S. (2008). Crisis at Colombia border spills into diplomatic realm. *New York Times*, *4*.

Santamaria, J. (2020). When a stranger shall sojourn with thee': The impact of the Venezuelan exodus on Colombian labor markets. *University of Minnesota, Department of Applied Economics, working paper*, *51422*.

Schneeweis, N. (2015). Immigrant concentration in schools: Consequences for native and migrant students. *Labour Economics*, *35*, 63–76.

Seah, K. (2021). Impact of immigrant peers on native students' achievement: Evidence from Australia, Canada and the United States. *Australian Economic Review*, *54*(1), 94–116.

Sequera, V. (2018). Venezuelans report big weight losses in 2017 as hunger hits. *World Hunger News*.

The Guardian. (2017, May). Venezuela's infant mortality, maternal mortality and malaria cases soar. *The Guardian Article*. Retrieved from https://www.theguardian.com/global-development/2017/may/09/venezuela-public-health-crisis-infant-mortality-maternal-malaria

Tonello, M. (2016). Peer effects of non-native students on natives' educational outcomes: Mechanisms and evidence. *Empirical Economics*, *51*, 383–414.

Tumen, S. (2021). The effect of refugees on native adolescents' test scores: Quasi-experimental evidence from PISA. *Journal of Development Economics*, *150*, 102633.

UNHCR. (2021). *Global trends: Forced displacement in 2021* (Tech. Rep.). Technical Report, The UN Refugee Agency.

van der Werf, C. (2021). *The impact of refugees on native students' educational outcomes.* Retrieved from https://www.cynthiavanderwerf.com/uploads/4/3/5/0/43507971/van_der_werf_2021.pdf

Appendix A Additional figures and tables

A.1 Other Papers Information

Table A.1: Country of destiny and origin of immigrants by paper

Paper	Natives country	Immigrants country	Income level classification	GDP per capita ($)
Ballatore et al. (2018)	Italy	Not specified	High income	42,562.5
Bossavie (2020)	Netherlands	Not specified	High income	57,334.7
Contini (2013)	Italy	Albania, China, Morocco, among others	High income	42,562.5
Contreras and Gallardo (2022)	Chile	Haiti, Venezuela	High income	25,412.7
Entorf and Lauk (2008)	Europe, Canada, Australia, New Zealand	Not specified	High income	
Figlio and Özek (2019)	USA	Haiti	High income	63,635.8
Figlio et al. (2021)	USA	Not specified	High income	63,635.8
Fletcher et al. (2021)	USA	Not specified	High income	63,635.8
Frattini and Meschi (2019)	Italy	Albania, India, Peru, among others	High income	42,562.5
Geay et al. (2013)	UK	Bangladesh, China, India, among others	High income	45,567.6
Gould et al. (2009)	Israel	URSS	High income	42,379.1
Green et al. (2022)	Norway	English speakers	High income	65,915.5
Hermansen and Birkelund (2015)	Norway	Pakistan, Turkey, Vietnam, among others	High income	65,915.5
Hunt (2017)	USA	Mainly Hispanic	High income	63,635.8
Jensen and Rasmussen (2011)	Denmark	Mainly non-Western Europeans	High income	58,802.9
Luksic (2022)	Chile	Venezuela, Haiti, Bolivia, among others	High income	25,412.7
Martínez and Martínez (2023)	Peru	Venezuela	Upper middle income	12,533.8
Morales (2022)	USA	Not specified	High income	63,635.8
Ohinata and Van Ours (2013)	Netherlands	Not specified	High income	57,334.7
Ohinata and Van Ours (2016)	Netherlands	Turkey, Morocco, Vietnam, among others	High income	57,334.7
Pedraja-Chaparro et al. (2016)	Spain	Not specified	High income	38,319.3
Schneeweis (2015)	Austria	Yugoslavia, Turkey, URSS	High income	53,817.3
Seah (2021)	Australia, Canada, USA	Not specified	High income	49,609.2
Tonello (2016)	Italy	Not specified	High income	42,562.5
Tumen (2021)	Turkey	Syria	Upper middle income	31,772.1
van der Werf (2021)	USA	South East Asia	High income	63,635.8

Note: "Natives" and "Immigrants" columns presents the country of destiny and origin of immigrants for every paper, respectively. Information of columns "Income level classification" and " GDP per capita" are retrieved from World Bank databases and are related to the destination country in "Natives" column. GDP per capita is PPP constant to 2017 international $us.

A.2 Timeline

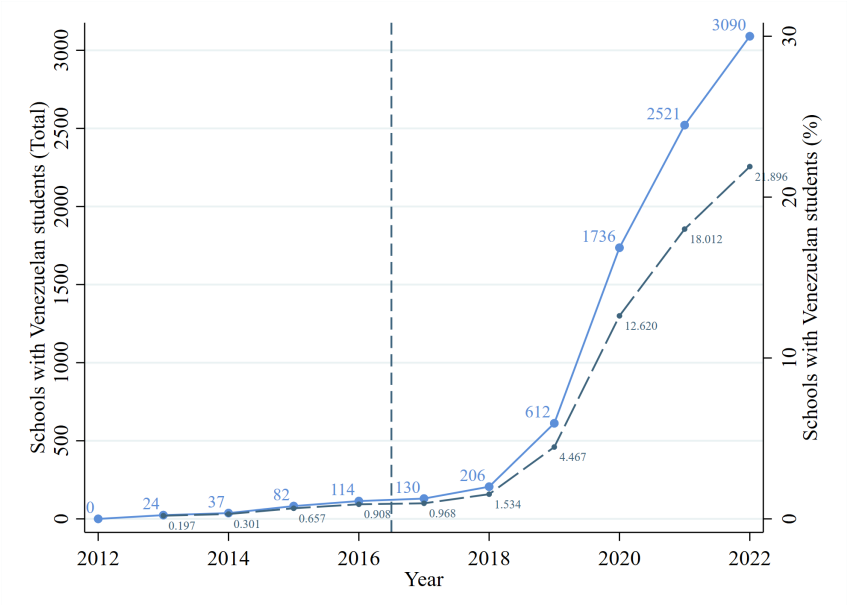
Table A.2: Timeline of events in Venezuela and Colombia

December 1998	Hugo Chávez's first presidential election.
September 2003	Beginning of the international oil "boom".
March 2013	Death of Hugo Chávez.
April 2013	Nicolás Maduro's first presidential election.
September 2014	End of the international oil "boom.
August 2015	Closure of borders between Colombia and Venezuela.
August 2016	Total opening of borders between Colombia and Venezuela.
May 2018	Second presidential election of Nicolás Maduro.

Note: This Table summarizes the important events between Colombia and Venezuela that developed the Venezuelan exodus. The information was verified in news and official sources of the Colombian and Venezuelan governments.

A.3 Schools with Venezuelan students

Figure A.1: Schools with Venezuelan students taking Saber 11 tests (2012-2022)



Note: The light blue line, with a solid line, measured on the left axis, corresponds to the total number of schools with at least one Venezuelan students taking Saber 11 tests each year. The dark blue line, with long dashes, measured on the right axis, corresponds to the percentage of schools with at least one Venezuelan students taking Saber 11 tests each year. The dark blue vertical line, with dashes, corresponds to the opening of borders between Colombia and Venezuela in mid-2016.
Source: Own elaboration using Saber 11 tests from 2012 to 2022.

A.4 Descriptive statistics by schools

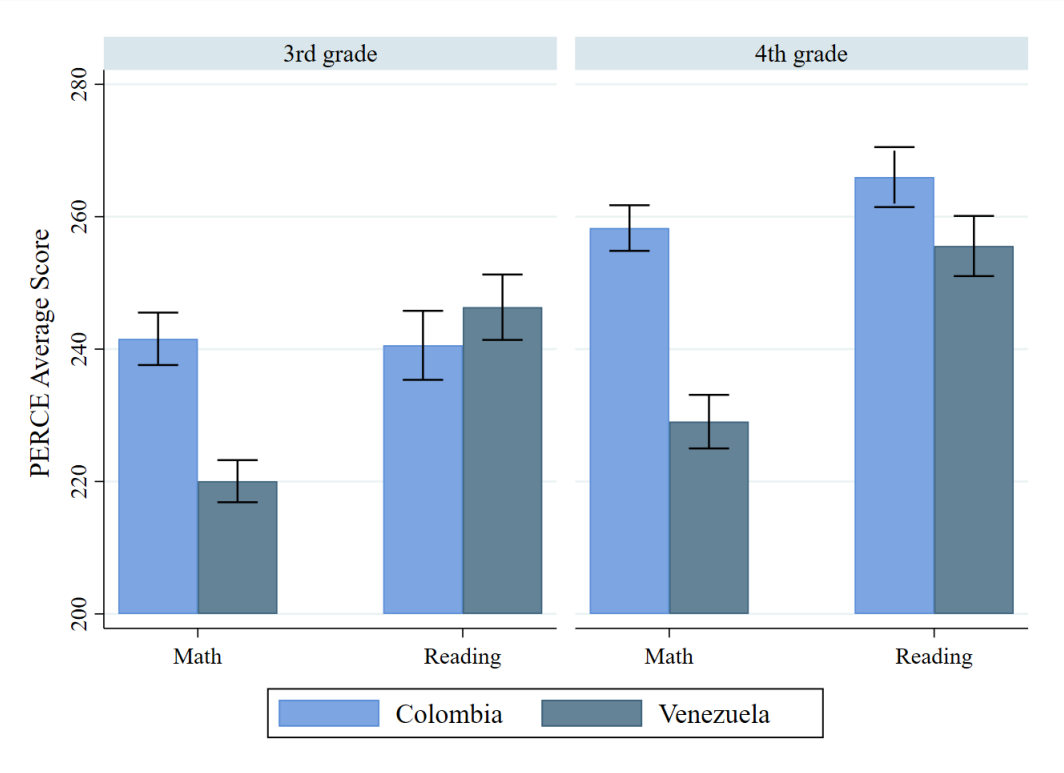
Table A.3: Characteristics of schools with and without Venezuelan immigrants (2013-2022)

Variable	Schools without immigrants		Schools with immigrants		
	Average	Standard Dev	Average	Standard Dev	Difference
Math	−0.22	0.61	−0.12	0.56	−0.10***
Reading	−0.22	0.60	−0.10	0.54	−0.12***
Summary	−0.24	0.67	−0.13	0.64	−0.11***
Natural Sciences	−0.20	0.60	−0.12	0.56	−0.08***
Social Sciences	−0.19	0.57	−0.11	0.53	−0.08***
English	−0.20	0.64	−0.10	0.66	−0.10***
Female (%)	0.53	0.15	0.54	0.11	0.00**
Age	17.78	2.04	17.43	1.26	0.35***
Rural school (%)	0.36	0.48	0.16	0.37	0.20***
Low mother education (%)	0.55	0.28	0.47	0.21	0.08***
Middle mother education (%)	0.27	0.15	0.33	0.11	−0.06***
High mother education (%)	0.17	0.21	0.20	0.19	−0.02***
Low father education (%)	0.62	0.27	0.54	0.21	0.08***
Middle father education (%)	0.24	0.15	0.30	0.11	−0.06***
High father education (%)	0.15	0.19	0.17	0.18	−0.02***
Stratum 1 (%)	0.44	0.31	0.36	0.27	0.08***
Stratum 2 (%)	0.32	0.22	0.36	0.21	−0.04***
Stratum 3 (%)	0.16	0.19	0.19	0.19	−0.03***
Stratum 4 (%)	0.05	0.09	0.05	0.09	0.00***
Stratum 5 (%)	0.02	0.05	0.02	0.06	0.00***
Stratum 6 (%)	0.01	0.04	0.01	0.06	0.00***
Calendar A (%)	0.97	0.16	0.97	0.15	−0.01***
Schools	15,152		4,950		20,102

Note: The statistics presented are the averages of the averages of each school taking into account only the natives who reported that they were born in Colombia. Math, Reading, Natural Sciences, Social Sciences and English correspond to the standardized scores obtained in the Saber 11 tests. Summary is the average standardized score of the five tests. The Strata i variables correspond to the proportion of students living in a home in strata i. The low mother education refers to incomplete secondary school; the middle category includes complete secondary and incomplete superior; and high category includes full superior or more.
Source: Own elaboration using Saber 11 tests from 2013 to 2022.

A.5 Academic achievements by country

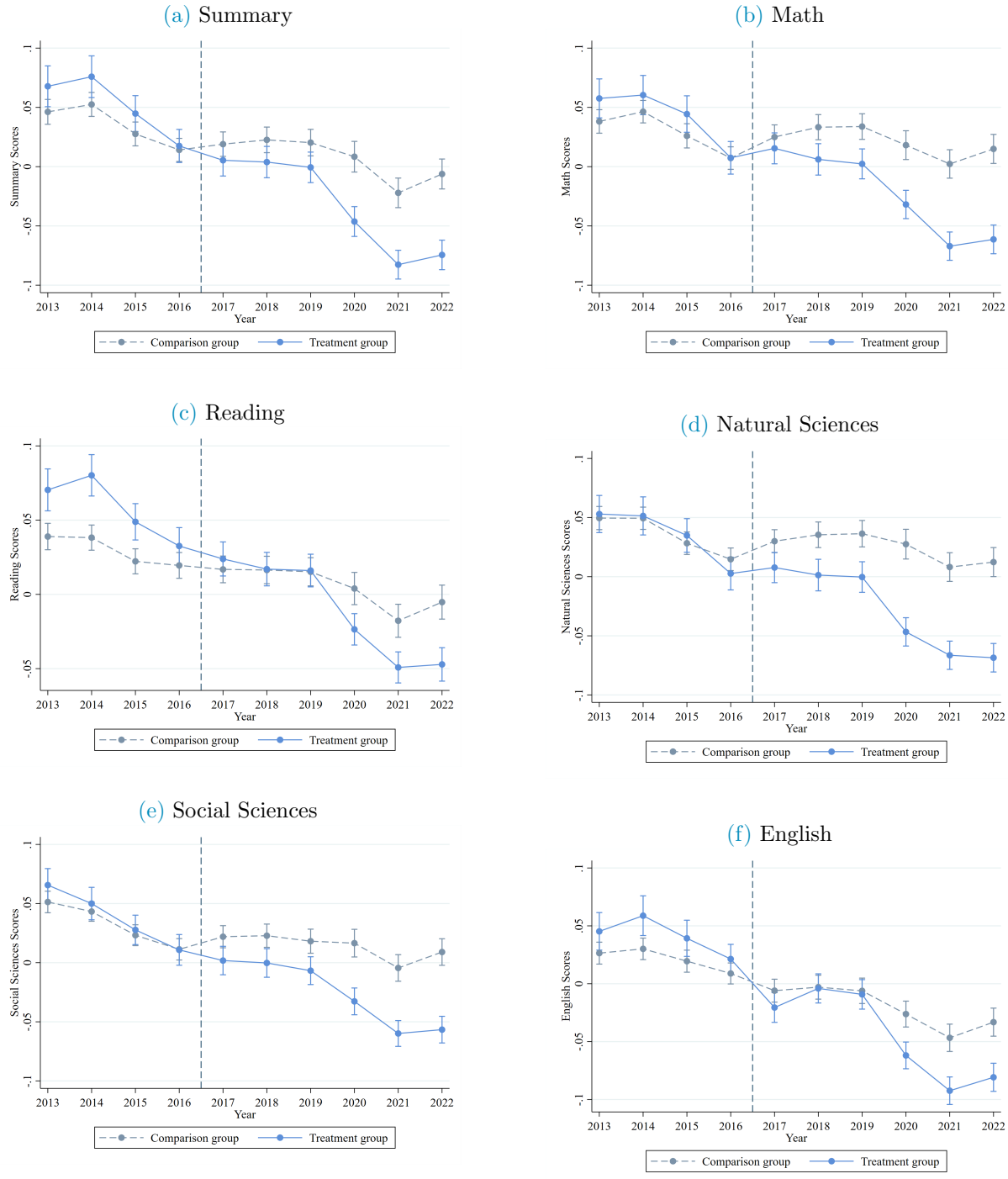
Figure A.2: Average scores of Colombia and Venezuela in PERCE (1997)



Note: 95% confidence intervals for each estimated average score are presented. The Figure shows the average scores achieved by Colombian (light blue bar) and Venezuelan (dark blue bar) students in international Math and Reading tests. The sample size for the study was 4,305 Colombian students and 3,691 Venezuelan students. The study was conducted by UNESCO's Latin American Laboratory for the Assessment of the Quality of Education.
Source: Own elaboration using data from the First Regional Comparative and Explanatory Study (PERCE) of 1997.

A.6 Trends of the scores for treated and comparison students

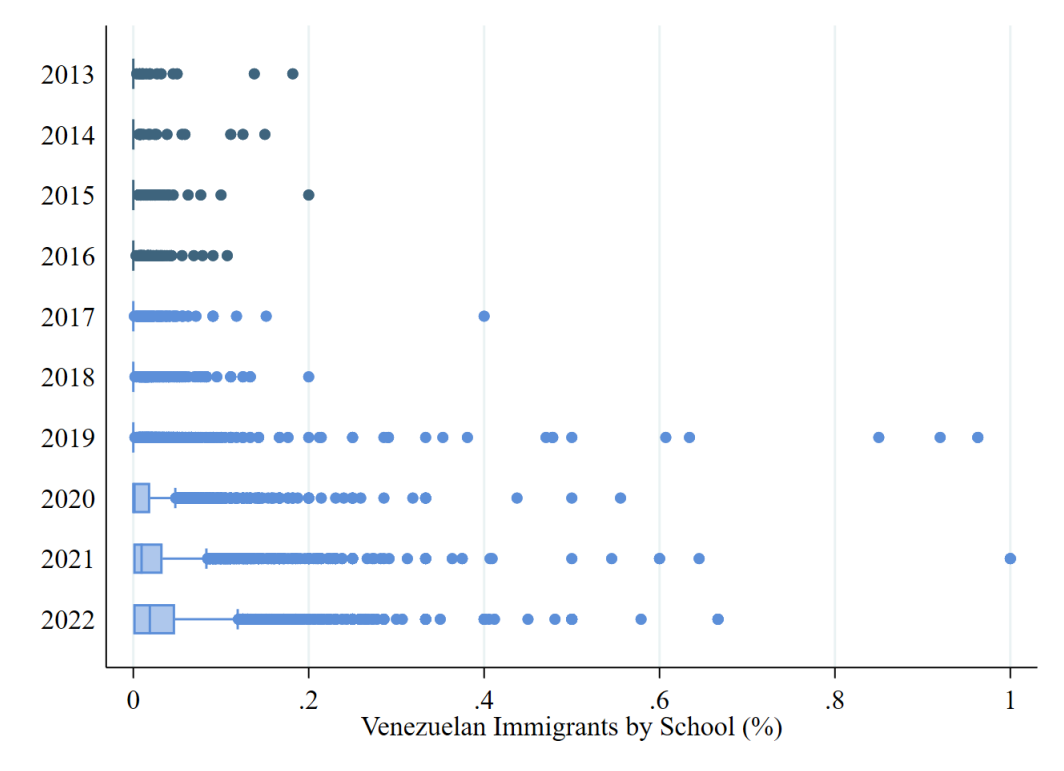
Figure A.3: Trends of the scores for treated and comparison students



Note: 95% confidence intervals are presented. Robust standard errors are clustered at the school level. Each panel of the figure presents the trends for the average standardized scores for both, treatment and comparison groups. Treatment group are the students in the treated departments (Norte de Santander and La Guajira) and the comparison group are the students in the remaining departments.
Source: Own elaboration using Saber 11 tests from 2013 to 2022.

A.7 Share of immigrants in treated schools by year

Figure A.4: Proportion of Venezuelan immigrants by school (%, 2013-2022)



Note: Box plot by year of the share of Venezuelan immigrants for treated schools by year are presented. Dark blue distribution corresponds to the period before the opening of borders and the light blue to the period after the opening of borders.
Source: Own elaboration using Saber 11 tests from 2013 to 2022.

A.8 Main results

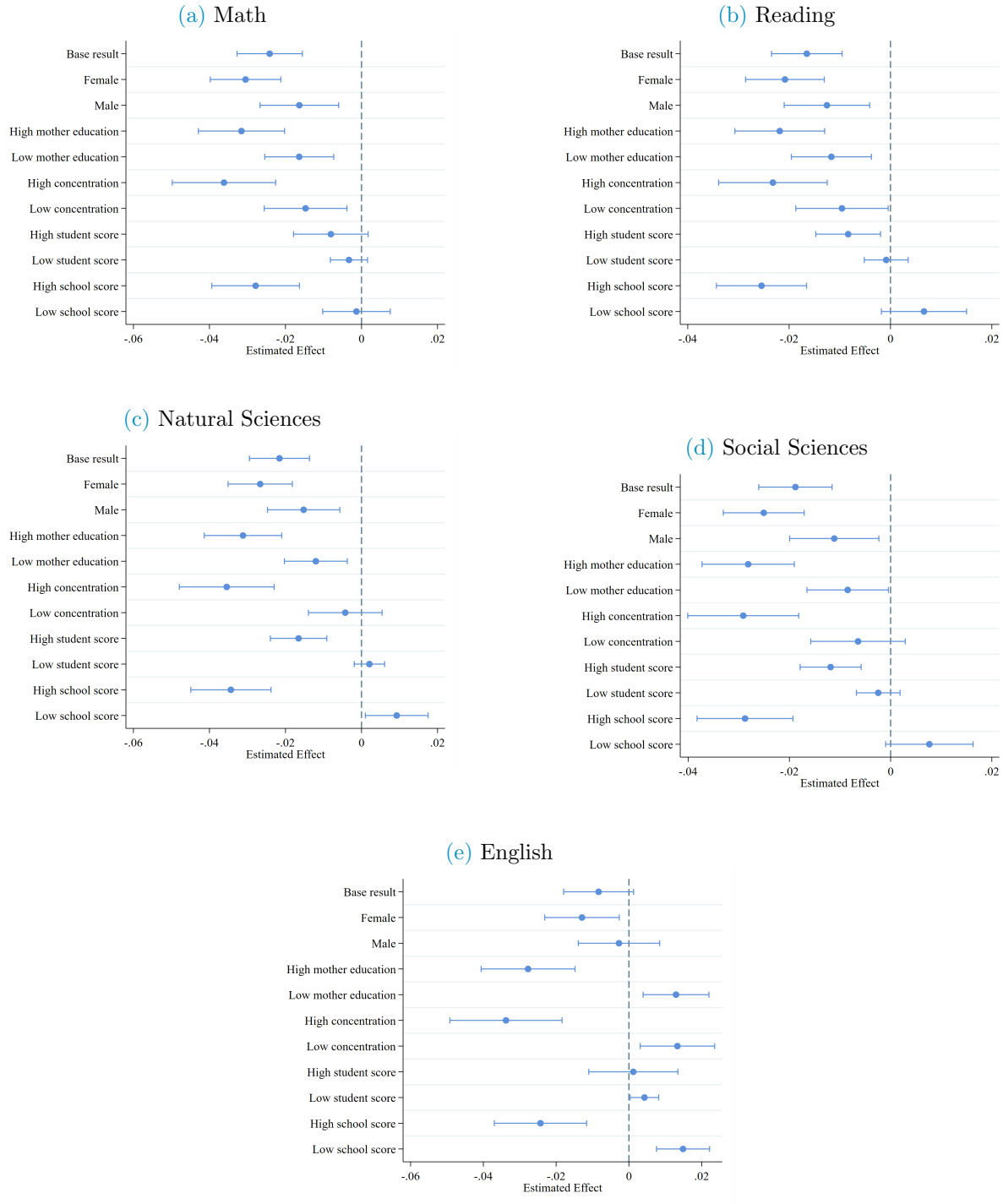
Table A.4: Main results for natives including all estimated effects

Variables	Summary (1)	Math (2)	Reading (3)	Natural Sci (4)	Social Sci (5)	English (6)
β	-0.018***	-0.024***	-0.017***	-0.022***	-0.019***	-0.008*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
Female	-0.175***	-0.297***	-0.039***	-0.217***	-0.118***	-0.087***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age	-0.309***	-0.316***	-0.262***	-0.302***	-0.259***	-0.211***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Age squared	0.006***	0.006***	0.005***	0.006***	0.005***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Rural area	-0.025	-0.048*	-0.034	-0.013	-0.022	0.007
	(0.023)	(0.025)	(0.022)	(0.030)	(0.022)	(0.019)
Middle mother education	0.074***	0.066***	0.074***	0.063***	0.059***	0.063***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
High mother education	0.218***	0.176***	0.201***	0.186***	0.196***	0.197***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Middle father education	0.074***	0.053***	0.075***	0.057***	0.068***	0.070***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
High father education	0.210***	0.155***	0.186***	0.176***	0.192***	0.207***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Total students	-0.050***	-0.046***	-0.037***	-0.049***	-0.048***	-0.035***
	(0.010)	(0.011)	(0.008)	(0.010)	(0.010)	(0.007)
% Female	0.062***	0.068***	0.047***	0.070***	0.061***	0.035***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Students	4,541,196	4,541,196	4,541,196	4,541,196	4,541,196	4,541,196
Schools	19,498	19,498	19,498	19,498	19,498	19,498
Family Strata	Yes	Yes	Yes	Yes	Yes	Yes
School Day	Yes	Yes	Yes	Yes	Yes	Yes
Year & Dept FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: *** ; ** ; * significant at the 1%, 5%, and 10% level, respectively. Robust standard errors by school clusters in parentheses. The natives are the students who reported that they were born in Colombia. Summary is the average standardized score between Math, Reading, Natural Sciences, Social Sciences and English. The reported coefficients that correspond to β are the same as in Table 2. Estimates in all columns include year and department fixed effects and student- and school-level controls.
Source: Own elaboration using Saber 11 tests from 2013 to 2022.

A.9 Additional heterogeneous effects

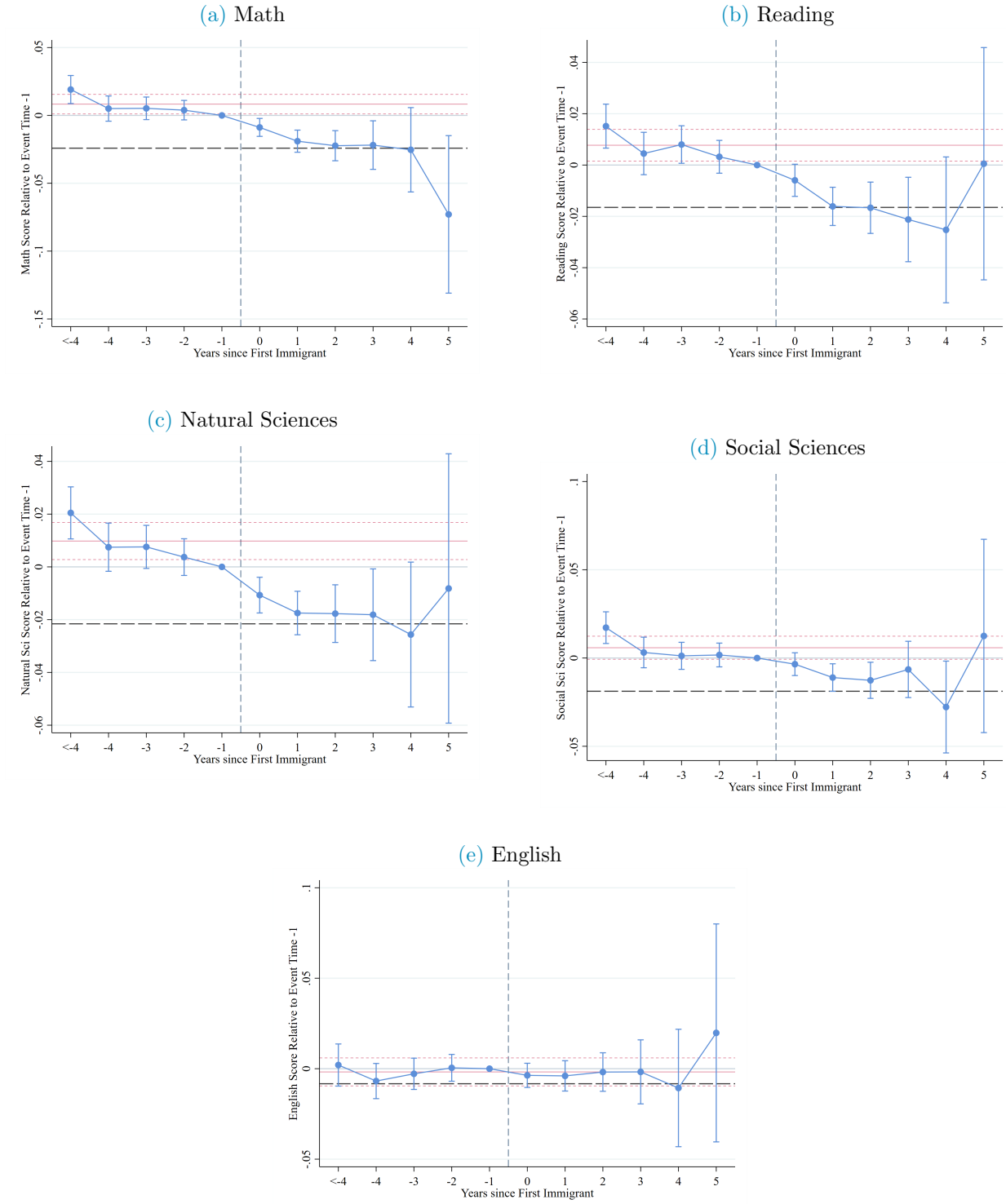
Figure A.5: Heterogeneous effects by different groups of native students



Note: 95% confidence intervals are presented. Robust standard errors are clustered at the school level. The reported coefficients are those that correspond to β from equation 1 for each of the groups. The effects of *High mother education* and *Low mother education* are estimates for natives who have mothers with high and medium educational levels and for natives with mothers with low education, respectively. The effects of *High concentration* are for natives who attend schools with more than 50% of mothers with high and middle education and *Low concentration* are for natives in schools with less than 50% of mothers with high and middle education. The *High student score* and *Low student score* estimates are for native students who obtain scores above and below the median year by year, respectively. The coefficients for *High school score* and *Low school score* are for native students attending schools with average scores above and below the school median year-to-year, respectively.
Source: Own elaboration using Saber 11 tests from 2013 to 2022.

A.10 Additional event study results

Figure A.6: Event study of the effects of immigration on learning of native students



Note: 95% confidence intervals are presented. Robust standard errors are clustered at the school level. The reported coefficients are those that correspond to β_{-j} and β_j from equation 4. The dark blue vertical line, with dashes, corresponds to the moment of treatment of schools. The dark blue horizontal line, with dashes, shows the results obtained in column 3 of Table 2 for the summary score.
Source: Own elaboration using Saber 11 tests from 2013 to 2022.

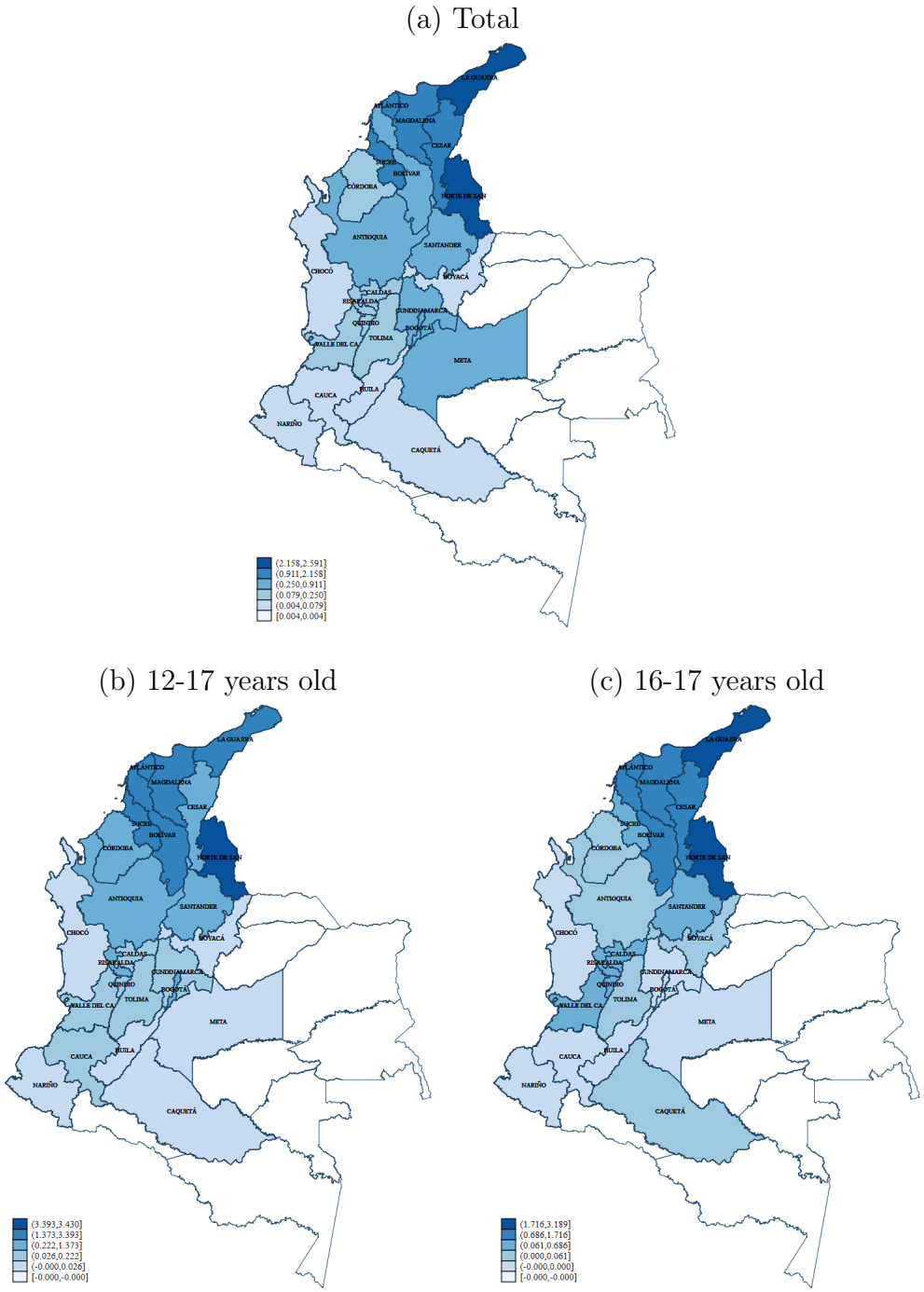
Table A.5: Event study coefficients of the effects of immigration on learning of native students

Variables	Summary (1)	Math (2)	Reading (3)	Natural Sci (4)	Social Sci (5)	English (6)
Relative to Event Time < -4	0.014***	0.019***	0.015***	0.020***	0.017***	0.002
	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.006)
Relative to Event Time -4	0.004	0.005	0.004	0.007	0.003	-0.007
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
Relative to Event Time -3	0.006	0.005	0.008**	0.008*	0.001	-0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Relative to Event Time -2	0.003	0.004	0.003	0.004	0.002	0.000
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
Relative to Event Time 0	-0.007**	-0.009***	-0.006*	-0.011***	-0.003	-0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Relative to Event Time 1	-0.014***	-0.019***	-0.016***	-0.017***	-0.011***	-0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Relative to Event Time 2	-0.015***	-0.022***	-0.017***	-0.018***	-0.013**	-0.002
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.005)
Relative to Event Time 3	-0.014	-0.022**	-0.021**	-0.018**	-0.006	-0.002
	(0.008)	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)
Relative to Event Time 4	-0.030**	-0.025	-0.025*	-0.026*	-0.028**	-0.011
	(0.015)	(0.016)	(0.014)	(0.014)	(0.013)	(0.017)
Relative to Event Time 5	-0.025	-0.073**	0.001	-0.008	0.013	0.020
	(0.028)	(0.030)	(0.023)	(0.026)	(0.028)	(0.031)
DID Effect	-0.018***	-0.024***	-0.017***	-0.022***	-0.019***	-0.008*
$H0$ joint pre-treatment	[0.007]*	[0.008]**	[0.008]**	[0.010]***	[0.006]*	[-0.002]
Students	4,541,196	4,541,196	4,541,196	4,541,196	4,541,196	4,541,196
Schools	19,498	19,498	19,498	19,498	19,498	19,498
FE & Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: *** ; ** ; * significant at the 1%, 5%, and 10% level, respectively. Robust standard errors by school clusters in parentheses. The natives are the students who reported that they were born in Colombia. Summary is the average standardized score between Math, Reading, Natural Sciences, Social Sciences and English. The reported coefficients that correspond to β are the same as in Table 2. Estimates in all columns include year and department fixed effects and student- and school-level controls.
Source: Own elaboration using Saber 11 tests from 2013 to 2022.

A.11 Venezuelan immigrants by department - Population variations

Figure A.7: Venezuelan immigrants by department (%, 2017)



Note: Venezuelan immigrants are considered those individuals who reported being born in Venezuela in the responses to the Migration Module. The departments with white color, which do not have the corresponding name, do not have observations in the GEIH. Most of the departments without observations are located in the Amazon region and have a low population density. The sampling weights from the survey carried out in the month of August were used to calculate the departmental proportions.
Source: Own elaboration using GEIH Migration Module of 2017.

A.12 Scores by departments

Table A.6: Scores by departments and by natives or immigrants (2013-2022)

Variable	Natives		Immigrants		
	Average	Standard Dev	Average	Standard Dev	Difference
Panel A: Control departments					
Summary	0.04	1.00	−0.08	0.97	0.12***
Math	0.04	0.99	−0.15	0.97	0.19***
Reading	0.04	0.99	−0.04	0.95	0.08***
Natural Sci	0.04	0.99	−0.08	0.96	0.12***
Social Sci	0.04	1.00	−0.06	0.94	0.10***
English	0.04	1.00	−0.02	1.05	0.05***
Students	4,472,547		11,262		4,483,809
Panel B: Treated departments					
Summary	−0.07	0.96	−0.06	0.94	−0.01
Math	0.00	0.99	−0.07	0.99	0.08***
Reading	−0.07	0.98	0.00	0.95	−0.07**
Natural Sci	−0.03	0.98	−0.03	0.94	0.00
Social Sci	−0.08	0.98	−0.03	0.93	−0.05**
English	−0.13	0.89	−0.11	0.92	−0.01
Students	200,785		1,385		202,170

Note: The natives are the students who reported that they were born in Colombia and the immigrants are those who were born in Venezuela. Math, Reading, Natural Sciences, Social Sciences and English correspond to the standardized scores obtained in the Saber 11 tests. Summary is the average standardized score of the five tests. Panel A statistics are for comparison departments and Panel B are for treated departments (*Norte de Santander* and *La Guajira*).
Source: Own elaboration using Saber 11 tests from 2013 to 2022.