Diversity in Schools:

Immigrants and the Educational Performance of U.S.-Born Students[1]

David Figlio, University of Rochester and NBER

Paola Giuliano, UCLA and NBER

Riccardo Marchingiglio, Analysis Group

Umut Ozek, RAND Corporation

Paola Sapienza, Northwestern University and NBER

Current Draft: January 2023

Abstract

We study the effect of exposure to immigrants on the educational outcomes of U.S.-born students, using a unique dataset combining population-level birth and school records from Florida. This research question is complicated by substantial school selection of U.S.-born students, especially among White and comparatively affluent students, in response to the presence of immigrant students in the school. We propose a new identification strategy, comparing sibling outcomes with the inclusion of family fixed effects, to partial out the unobserved non-random selection of native-born families into schools. We find that the presence of immigrant students has a positive effect on the academic achievement of U.S.-born students, especially for students from disadvantaged backgrounds. Moreover, the presence of immigrants does not negatively affect the performance of affluent U.S.-born students, who typically show a higher academic achievement compared to immigrant students. We provide suggestive evidence on potential channels.

JEL Classification: JEL No. I21, I24, J15
Keywords: Immigrant students; Educational attainment.

[1] We thank the Editor, Thomas Chaney, and three anonymous referees for comments that greatly improved the paper. We also thank Jennifer Hunt, Andrea Ichino, Victor Lavy, Daniele Paserman, Bruce Sacerdote, Cynthia van der Werf, as well as participants at various seminars and conferences. For truly outstanding research assistance, we thank Gaia Dossi and Matteo Giugovaz. We are especially grateful to the Florida Departments of Education and Health for providing the linked anonymous population-level administrative data that permitted this analysis to take place. All errors and opinions are those of the authors and do not reflect those of the funders or the Florida Departments of Education and Health. Interested researchers may request access to Florida administrative data through the processes described at fldoe.org/accountability/accountability-reporting/external-research-requests. This research was funded with support from the Russell Sage Foundation. Riccardo Marchingiglio performed this research while affiliated with the Department of Economics at Northwestern University. The views presented are those of the authors and do not necessarily reflect those of Analysis Group, Inc. or its clients.

1. Introduction

Over the past 50 years, immigration into the United States has risen dramatically. As a result, several public-school districts have a high concentration of foreign-born students (as high as 42% in Miami-Dade County, Florida[2]). These trends have generated a policy debate about the effects of immigration on public education and the perceived costs that immigrants may impose on public schools, local governments, and educational outcomes of the U.S.-born student population.

Given the sheer magnitude of the number of immigrants in U.S. schools and their unique cultural backgrounds, studying their impact on U.S.-born students is of first order importance. On one hand, immigrants may face challenges in assimilation that may require additional school resources which could be taken away from U.S.-born students (Fix and Zimmerman, 1993). On the other hand, some groups of immigrants tend to outperform non-immigrant students with similar socio-economic backgrounds, perhaps due to difficult-to-measure attributes such as hard work and resilience. The literature has shown that, after controlling for socio-economic status and academic ability, the cultural orientation of immigrant families affects students' beliefs and expectations and their academic success. Hsin and Xie (2014) show that Asian and Asian-American's greater academic success is linked to cultural differences in beliefs regarding the connection between effort and achievement, while Figlio et al. (2019) relate the overperformance of immigrants from specific countries to cultures that emphasize deferred gratification and self-control. Exposure to students with greater work ethic can in turn positively affect U.S.-born students' attitudes and behavior.

The academic research about the impact of immigrant students on the educational performance of U.S.-born students is limited, primarily due to two important empirical challenges.[3] First, immigrant students are not randomly assigned to schools, and are more likely to enroll in schools educating students from disadvantaged backgrounds (Card, 2001). Second, U.S.-born students, especially those from comparatively affluent families, may decide to leave when a large share of immigrant students move into their school district. Indeed, evidence shows that in the US, following

[2] Authors' calculation from the American Community Survey 2015-19.
[3] There are a few studies that examine the effects of immigrant students on U.S.-born student outcomes. For example, Schwartz and Stiefel (2011) use within-school variation and find a negative effect of immigrant share on the performance of U.S.-born students in New York City public schools. McHenry (2015) and Hunt (2016) examine the effects on high school completion rates of native-born students and find positive immigrant effects, especially among students from disadvantaged backgrounds. Similarly, Neymotin (2009) finds no adverse effect of immigration on the SAT-scores and college application patterns of U.S.-born students. In a quite different context, some studies investigate the effects of refugees on U.S.-born student outcomes (Figlio and Ozek, 2019; Morales 2020; Ozek, forthcoming; Van der Werf 2021).

an influx of disadvantaged students and immigrants, affluent, especially White, students move to private schools or districts with higher socio-economic status (SES) families, a phenomenon which has been labeled "white flight" (Betts and Fairlie, 2003; Cascio and Lewis, 2012; Fairlie and Resch, 2002; Li, 2009). Both factors imply that immigrant exposure is negatively correlated with the SES of U.S.-born students. Therefore, research that does not address the non-random selection of U.S.-born students is likely to estimate a correlation between immigrant exposure and U.S.-born student outcomes that is more negative than the true relationship. The unique features of our data allow us to directly address both selection issues for the US.[4]

We study the effects of exposure to immigrants on the educational outcomes of U.S.-born students using administrative data from Florida that link population-level school records and birth records. There are several advantages in using this dataset. First, it is especially interesting to study this question in Florida because it is one of the largest immigrants receiving states in the U.S. with a lot of heterogeneity across school districts, in terms of number and type of immigrants.[5] Second, birth records allow us to identify *siblings* and control for all the observable and unobservable family characteristics (even family life-cycle characteristics) with the inclusion of family-year fixed effects. Third, the dataset follows individual students over time, thus allowing us to measure a cumulative exposure to immigrants. Using this information, our first identification strategy compares test scores in math and reading of siblings who experience different cumulative exposures to school-cohort-specific immigrant concentrations, holding the heterogeneity in family life cycle fixed. Further, because we have information on the entire population of students attending public schools during this period, we can employ a second identification strategy, using an instrumental variable approach, to address the possibility that families select schools differentially for each child. Specifically, we build a measure of predicted immigrant exposure using aggregate school-to-school transition probabilities, for each

---

[4] In contexts different from the US, the literature has found zero or negative effects (Jensen and Rasmussen, 2011; Brunello and Rocco, 2013; Ballatore et al., 2018; Tornello, 2016; Ohinata and van Ours 2013; Geay et al., 2013; and Schneeweis, 2015; Bossavie, 2020). Gould et al. (2009) successfully addressed the selection of immigrants into schools by exploiting an exogenous inflow of refugees from the Soviet Union that occurred in Israel during the 1990s. They find a negative effect of immigration on the probability of passing the high-school matriculation exam, affecting mostly poor Israelis.

[5] Florida has over four million foreign-born individuals (roughly corresponding to 20 percent of the population). Florida's foreign-born population is also diverse. While the foreign-born population is disproportionately Hispanic (including 23% Cubans and 7% Mexicans), 21% of immigrants come from non-Hispanic Caribbean countries, 11% from Asian countries, 10% from European countries, and 2% from African countries. The heterogeneity by country of origin of foreign-born residents in Florida is much greater than in Texas and California, where most foreign-born residents come from a single country, Mexico.

kid at each subsequent grade, starting from the first grade the student is observed. Two siblings who started in the same school (in different years) will have the same predicted transition matrix but a different predicted exposure to immigrants, which depends on their specific cohort. We also employ an alternative IV strategy that is robust to the situations in which families enroll their children in different initial schools.

Our empirical analysis proceeds in several steps. We first calculate a measure of cumulative immigrant exposure using the longitudinal aspect of our data. We then estimate a specification similar to the one commonly used in the extant literature in order to address the first form of nonrandom selection described above. This specification compares students with different exposures to immigrants, only controlling for school and grade fixed effects (both interacted with calendar year dummies). When we estimate this equation, we find a significant--although small in magnitude--negative correlation between the share of immigrants and the natives' scholastic performance in both mathematics and reading. But this specification does not address the non-random selection of U.S.-born students based on the expected immigrant concentration of the school. We therefore compare *siblings'* outcomes with the inclusion of family fixed effects. When we employ this specification, the estimated relationship between immigrant concentration and student outcomes becomes positive. This fundamental finding is unaffected by still more stringent identification strategies, such as including family-year fixed effects to control for family lifecycle changes, or instrumental variable approaches. The reason for the discrepancy between our findings and those that do not address nonrandom U.S.-born student selection into schools is that in the U.S. there is evidence that students indeed sort into schools based on immigrant concentration. We find that this sorting is concentrated among White and affluent students, consistent with the white flight literature (Betts and Fairlie, 2003; Cascio and Lewis, 2012). By contrast, our evidence suggests that, on average, Black and lower-SES students do not move away from schools or districts with a larger fraction of immigrants.

For the overall sample, the magnitude of the results indicates that moving from the tenth to the 90th percentile in the distribution of cumulative exposure to foreign-born students (1 percent and 13 percent, respectively) increases the score in mathematics and reading by 2.8 percent and 1.7 percent of a standard deviation, respectively. This effect corresponds to 8.5 percent of the differences in scores between children whose mother has a high school diploma and children whose mother has not completed high school. We also find that this effect is twice as large for free or reduced-price lunch (FRPL) eligible students and for Black students.

To investigate potential mechanisms and to aid interpretation of our findings, we first study whether the presence of immigrants is a proxy for other school or classroom characteristics. Specifically, we find that the correlation between the presence of immigrants and U.S.-born students' academic achievement is not explained by the demographic and SES school-cohort composition, level of school resources, diversity of the school body, or class segregation. These findings suggest that our main results are likely to be driven by specific traits, behavior, or attitudes of immigrants.

As a last step, we suggest that the behavior of the immigrants may have spillovers on their classmates either because they set examples with their peers (through their performance or behavior) or because they make the classroom environment more conducive to learning. To partially investigate this hypothesis, we conduct several additional analyses. First, we study whether the absolute performance of the immigrant students could explain our results. The reflection problem (Manski, 1993) and endogeneity issues do not allow the identification of the causal impact of the achievement (or behavior) of immigrants on the performance of U.S.-born students. Instead of including actual immigrant performance in the regression, we calculate a proxy for expected performance, by using the average immigrant academic performance by country of origin and multiply it by the fraction of immigrants in each grade/school/year. This measure of immigrant exposure weighted by country-of-origin average performance is a proxy for the potential academic achievement of immigrants. When added to our baseline specification, we show that the presence of immigrants with higher expected academic performance correlates with better scores of U.S.-born students in the overall sample and across every subsample. This channel does not affect the sign and magnitude of our main variable, the direct impact of immigrant exposure. Lazear (2001) suggests that student behavior can have spillovers in the classroom by facilitating learning through fewer disciplinary incidents. Adding a proxy for expected behavior of immigrants by country of origin to our baseline, we find that exposure to better-behaved immigrants is associated with better academic outcomes, albeit by a small amount. As before, our main variable of interest is unaffected.

Next, we investigate whether better *relative* performance of immigrants compared to U.S.-born students can explain our results. First, we document that immigrants underperform US-born affluent students and overperform low SES US-born students. Second, when immigrant students perform worse than their classmates, they do not have a negative spillover on them. This combined evidence suggests that our results could be driven by specific immigrant characteristics such as cultural attitudes. Following Figlio et al. (2019), we investigate whether exposure to immigrants coming from cultures that emphasize deferred gratification and self-control correlates with US-born academic achievements.

Our evidence supports this hypothesis, suggesting that horizontal transmission of attitudes and beliefs may play an important role in explaining educational performance.

2. Data and Variables of Interest

2.1 Data Sources

We use a dataset of school records for the state of Florida, maintained by the Florida Department of Education (FLDOE), merged with birth vital records from the Florida Bureau of Vital Statistics. The individual-level administrative data from the FLDOE contains information on K-12 students who attended Florida public schools between 2002-2003 and 2011-2012. The data contains for each child the results of the Florida Comprehensive Assessment Test (FCAT) in reading and mathematics administered annually to all students in grades 3 through 10, as well as disciplinary incidents. The dataset also contains information about the country of origin of the child and the language spoken at home. Birth vital records contain a larger set of SES measures for children born in Florida (such as maternal education, marital status, and age of the mother when the child was born), normally not included in school records.[6] The match with birth certificates allows us to identify children belonging to the same family and to exploit within family variation. Since data from birth certificates are available only for children born in Florida between 1994 and 2002, we limit our analysis to these cohorts.

2.2 Definition of Immigrants

Our goal is to study the effect of immigrant exposure on the performance of U.S.-born students. We define as *immigrants* all students born in a foreign country (the information on the country of origin is in the school administrative records).[7] One inherent challenge involves how to treat students born in Puerto Rico. On the one hand, Puerto Rico-born children are U.S. citizens, and should be considered similarly in the analysis to children born in Texas or Massachusetts. On the other

[6] Birth certificates and school records were matched using first and last names, date of birth and social security numbers. The sample of birth records consists of 2,047,633 observations. Of these, 1,652,333 were present in Florida public school data. The match rate of 81 percent is consistent with the percentage of children who are born in Florida, reside there until school age, and attend public school, as calculated from the Census and the American Community survey for the corresponding years. See Figlio et al. (2014) for details about the nature and additional evidence on the quality of the birth-school data merge.

[7] One complication in our data is that some US citizens born abroad (most notably because of parents serving in the military) are recorded as "foreign-born" in the data. There is no perfect way to address this limitation, but we try to partially bound the effect by excluding observations from the four Florida counties (Bay, Brevard, Clay, and Okaloosa) with large military concentrations to gauge whether our results are sensitive to their inclusion; we will discuss this result in Section 3.

hand, Puerto Rico-born children may be "othered" in a manner such that incumbent families react to their presence as if they are immigrants. Because we do not have birth certificates for Puerto Rican students, we are unable to include them in the sample of U.S.-born students. Therefore, the best way to treat Puerto Rican-born students in this analysis is not obvious: we can either consider them "foreign" students when calculating the foreign exposure measure or we can exclude them from the calculation. We adopt both strategies, in order to gauge the degree to which this empirical decision affects our results. In the baseline regression, we treat students born in Puerto Rico as "immigrants" on the ground that they are culturally distinct from other US citizens. However, in a robustness analysis, we do not include them in the construction of the immigrant exposure variable and the results are unchanged from the baseline. We therefore conclude that our choice of treatment of Puerto Rican-born students as "immigrants" does not influence our findings.

The birth certificates provide information on whether the mother was born abroad. Thus, we could have added to the first-generation immigrant children born in the U.S. with parents born abroad (second generation immigrants). Because we do not have information on the immigrant status of the father, we do not follow this strategy (although in robustness analyses, we study whether exposure to children of foreign-born mothers affects our analysis).

2.3 Measure of Immigrant Exposure

We adopt a cumulative measure of exposure to foreign-born students (or, immigrant exposure), in which we aggregate the share of foreign-born students to whom a U.S.-born student has been exposed from kindergarten to the time of observation (measured at the school-grade-year cell level). Our baseline definition of cumulative exposure for each student *i,* in school *s*, grade *g,* and academic year *t,* is:

$$Immigrant\ Exposure_{isgt} = \frac{1}{g}\sum_{g' \leq g} Immigrant\ Share_{isg't} \qquad (1)$$

In a robustness analysis, we will allow for a more flexible specification, where the effect of exposure to immigrant students decays over time.

2.4 Outcome Variables

Our main outcomes of interest are the Florida Comprehensive Assessment Test (FCAT) scores in mathematics and reading from grade 3 to grade 10 (the first and last grade of statewide testing). Because Florida transitioned to an updated version of the test, called FCAT 2.0, in 2011 and to aid in interpretation, we standardize the statewide test scores to zero mean and unit variance at the grade-year level over the entire population of students.

2.5 Individual Controls

In our specification, we include as controls several demographic variables (age in months, gender, birth order fixed effects, and race/ethnicity dummies), a measure of low-income status (a dummy for whether the student is eligible to receive free or reduced-price lunch or attend a "provision 2" school, where such a large fraction of students are eligible that individual documentation is not collected, as almost all students are presumptively eligible), a measure for whether the student receives special education services, and dummies for maternal education (high school graduate, some college and four years of college or more, with the excluded group given by mothers who dropped out of high school).[8]

2.6 Definition of U.S.-born Students and Construction of the Sample of Interest

We define as *U.S.-born students* all students born in the U.S. who speak English at home. Given the large fraction of second-generation immigrant students, we believe that the language restriction is more likely to select students who fully identify as Americans. However, in robustness analysis we remove this language restriction. In the Florida Department of Education data, we have the full population of students going to Florida public schools during the period 2002-2012.[9] Given our identification strategy, in our analysis we select the sample for which (1) we have test scores and (2) we can link school records to birth certificates. We report descriptive statistics for this sample in Columns 1 to 3 of Table A1.A, in the Appendix. This sample contains 8,010,198 (7,490,949)

[8] The race/ethnicity variable is collected by the Florida Department of Education according to the following categories: Hispanic/Latino of any race (Hispanic for brevity), American Indian or Alaska Native (classified into "Others"), Asian, Black or African American (Black for brevity), Native Hawaiian or Other Pacific Islander (classified into "Others"), White, Two or more races (classified into "Others"). To qualify for free or reduced lunch, the family income must be respectively below 185 percent and 130 percent of the federal income poverty. Provision 2 schools establish claiming percentages and serve all meals at no charge for a 4-year period. For details, see http://www.fns.usda.gov/school-meals/provisions-1-2-and-3. Categories for special education include mentally handicapped, orthopedically, speech, language, or visually impaired, deaf, or hard of hearing. It also includes students with emotional or behavioral disabilities, with autistic spectrum disorder, and other forms of serious disabilities (such as students with traumatic brain injuries). Maternal education data are reported in birth vital records.

[9] To understand the differences between U.S.-born students going to private and public schools in Florida, in Table A2 of the On-line Appendix, we report the descriptive statistics of the two groups. Using Census 2000 data, we compare the population of immigrant students attending public schools in Florida (93 percent) with those of the U.S.-born (88 percent). U.S.-born students, on average, are exposed to immigrant children who have lower SES than themselves, independently from the school setting: the family income of U.S.-born students going to private (public) schools is $102,409 ($55,838), while the income of immigrant students going to private (public) schools is $86,163 ($43,526). The patterns are similar for 2010.

observations for reading (math) scores.[10] The U.S.-born students with a birth certificate in the Florida Department of Education data are slightly positively selected compared to all students attending Florida public schools (standardized math and reading scores are 0.044 and 0.052). As our most demanding specification makes use of family-year fixed effects, we further restrict this sample to student-year observations in families with at least two children in the Florida public school system in a given academic year. This sample consists of 1,789,450 student-year observations (columns 4 to 6, Table A1.A). When we restrict the sample to U.S.-born students speaking English at home (Columns 7 to 9), we obtain 6,341,333 observations. From this sample, restricting to observations in families with at least two children in school in a given academic year leads to 1,450,138 observations for reading scores and 1,347,286 for math scores (Columns 10 to 12 of Table A1.A). Our final sample has similar standardized test scores to the original sample with birth certificates: 0.05 for math and 0.034 for reading (Table 1).

2.7 Characteristics of Immigrants

Columns 1 to 3 of Table A1.B report the sample statistics for the immigrant students who go to school with the sample of US-born students described in Columns 1 to 3 of Table A1.A. Immigrant students' performance in math (-0.097) and reading (-0.206) is lower than the one of U.S.-born students (0.044 and 0.052). Immigrants are also poorer (68 percent are FRPL-eligible) than U.S.-born (54 percent) and vary significantly in terms of racial background, language ability, and academic performance. In terms of racial composition, the majority identifies as Hispanic (61 percent), while among U.S.-born students only 22 percent are Hispanic. Immigrants are also more exposed to other immigrants (18 percent compared to U.S.-born students who are exposed to 8 percent of immigrants). Consistent with evidence in other domains where immigrants tend to commit fewer crimes than non-immigrants (Nunn et al., 2018), immigrant students participate in fewer disciplinary incidents (0.121) than U.S.-born students (0.137).

In Columns 4 to 12 of Table A1.B, we report the statistics of immigrants corresponding to the U.S.-born students described in Columns 4-12 of Table A1.A to verify that our selection of U.S.-born students does not lead to a different composition of immigrants in schools. Restricting to the sample of U.S.-born students with siblings in school and to those speaking English at home does not

[10] The discrepancy between the number of reading and math observations is because Florida stopped testing high school students after 2009-10 school year in math (when they transitioned to FCAT 2.0). Therefore, we have reading scores for 9th and 10th graders in 2010-11 and 2011-12, but no math scores.

change the characteristics of the foreign-born students compared to the sample of Column 1 to 3 of Table A1.B.

2.8 U.S.-born Students' Exposure to Immigrant Students

In the sample used in our regressions, students have an average cumulative exposure to immigrant students of 6 percent, but there is substantial variation across school-cohorts. Figure 1 shows the distribution of the fraction of immigrants by institution, grade, and year. Most schools have a fraction of immigrants lower than 10 percent; however, there is a non-trivial number of schools with a fraction of immigrants larger than 20 percent. Figure 2 maps the geographical distribution of immigrants in our sample and shows that the largest fractions tend to be concentrated in the southern part of the state. Figures 3A and 3B map schools in our sample divided by top and bottom decile in the distribution of immigrants for the whole state and the Miami-Dade County school district. Although the largest concentration appears to be in Miami-Dade, substantial variation also exists elsewhere.

To understand whether exposure changes over time, in Figure 4 we plot the average contemporaneous exposure for U.S.-born students by academic year. The average exposure by academic year appears to be stable, suggesting that there is not an increase over time in cohorts of immigrants.

We then look at whether U.S.-born students with a different racial or socio-economic background experience exposure to a different share and composition of immigrants. We start by splitting the sample of U.S.-born students by ethnicity/race (Figure 5) and we observe substantial differences in contemporaneous exposure to foreign-born. U.S.-born White students experience the lowest exposure to immigrants (around 6 percent), U.S.-born Hispanic students the largest (around 12 percent), and U.S.-born Black students somewhere in between (8 percent).

Table 2 and 3 lists the top ten countries of origin and ethnicity of immigrants in Florida facing our sample of U.S.-born students and facing the sub-samples of U.S.-born divided by ethnicity/race.[11] The top ten countries of origin in the overall sample are all Latin American countries. Together they constitute 65 percent of the immigrant sample. For the school-specific cohorts where most U.S.-born students are White, Hispanic, or Black the results vary.[12] In school-cohorts where most U.S.-born

[11] Note that, as mentioned above, we consider models in which we treat Puerto Ricans, all of whom are US citizens, either as "U.S.-born" or as "immigrants." For the purposes of Table 2, we count Puerto Ricans as "immigrants," so that the reader can gauge the share of Puerto Ricans in the overall Florida student population.
[12] There are 4,158 schools across all years in our main sample, and 3,676 have at least one foreign-born student in one cohort. 61,836 school-specific cohorts out of 84,019 have at least one foreign-born student. Among the

students are White, Mexico represents the largest fraction (13 percent); in addition, several non-Latin American countries are at the top of the distribution: Germany (5 percent), Canada (4 percent) and China (3 percent). In school-cohorts where most U.S.-born students are Hispanic, immigrants come mostly from Latin American countries, especially Cuba (46 percent). The ten largest countries of origin represent 85 percent of the overall immigrant distribution. Finally, in school-specific cohorts where most U.S.-born students are Black, the largest fraction of immigrants comes from Haiti (41 percent) and Jamaica (13 percent), and 78 percent of the immigrant exposure comes from ten countries.

U.S.-born students are exposed to Hispanic immigrants (62 percent), followed by Black immigrants (17 percent), and White immigrants (13 percent). However, there is a large heterogeneity in exposure once we split the schools by predominant ethnicities/races, confirming high level of sorting. A long literature in demography and sociology demonstrates that the tendency for immigrants to disproportionately co-locate with native-born families of the same ethnic/racial background is not specific to Florida, but it is prevalent across the United States (Alba and Logan, 1993; Alba et al., 1999; Friedman and Rosenbaum, 2004; Logan and Alba, 1993; Rosenbaum et al., 1999; Rosenbaum and Friedman, 2001, 2004; Schill, Friedman, and Rosenbaum, 1998) and that ethnicity/race is a dominant factor in immigrant residential location (Freeman, 2002; Rosenbaum and Friedman, 2004). Even in the suburbs, immigrant families tend to settle near same ethnicity/race native-born families (Alba et al., 1999), and the patterns of racial similarity in immigrant residential location are similar for new immigrant destinations and established immigrant gateways alike, though there may be slightly less immigrant racial segregation in new immigrant destinations than in established gateways (Friedman et al., 2005; Park and Iceland, 2011). Therefore, there is little evidence that the ethnic/racial similarity between immigrants and native-born students observed in Florida is unusual by national U.S standards. Taken all together, these initial descriptive statistics show that U.S.-born students are exposed to different subgroups of immigrant students, depending on their ethnicity/race.

3. Empirical Analysis

3.1 Main Results

61,836 school-specific cohorts with at least one foreign-born student, 27,067 school-specific cohorts have a majority of White U.S.-born students, 8,336 school-specific cohorts have a majority of Black U.S.-born students, while 6,326 school-specific cohorts have a majority of Hispanic U.S.-born students. The remaining schools in our sample have either a foreign-born majority or a U.S.-born majority of another racial/ethnic group.

Tables 4 Panel A and B present our main results. We regress our outcomes of interest, standardized test scores in math and reading, Y_{isgt}, of a student *i*, attending school *s*, in grade *g*, during the academic year *t* on $Immigrant\ Exposure_{isgt}$ defined in equation (1).[13] Our most demanding specification is the following:

$$Y_{isgt} = \alpha + Immigrant\ Exposure_{isgt}\beta + Z'_{isgt}\gamma + \delta_{gt} + \vartheta_{st} + \phi_{ft} + \varepsilon_{isgt} \qquad (2)$$

where Z'_{isgt} is a vector of individual characteristics, including gender, age in months, whether the student is a special-education student, birth order fixed effects, race, and FRPL eligibility; δ_{gt} are grade-year fixed effects and ϑ_{st} are school-year fixed effects; ϕ_{ft} are family-year fixed effects. Below the coefficient, in round brackets, we report robust standard errors clustered by school-cohort;[14] in square brackets, we report the standardized beta coefficient which indicates how one standard deviation increase in immigrant exposure translates into a standard deviation change of the test scores in math and readings.

In Column 1, we start by running a specification only controlling for the non-linear interaction of grade-year fixed effects (δ_{gt}), school-year fixed effects (ϑ_{st}), and a limited set of individual controls, Z'_{isgt} (age in months, gender, birth order fixed effects, and whether the student has some special education needs). The results are consistent with the previous literature (Schwartz and Stiefel, 2011): a significant negative correlation between the share of immigrants and the natives' scholastic performance both in mathematics and reading. The beta coefficient of cumulative immigrant exposure for the math score regressions (-0.006) is smaller than the corresponding beta coefficient for the reading score regressions (-0.01).

In Column 2, to correct for possible selection, we introduce a specific measure of students' SES (whether the student is FRPL-eligible) and control for ethnicity/race. The correlation between standardized test scores and fraction of immigrants becomes positive, albeit insignificant, for math, and remains negative, but insignificant for reading. In Column 3 we add, as an additional proxy for SES, maternal education (because this variable is missing for some observations, the number of

[13] The math and reading scores are standardized using the entire population of students. To make sure that the results are not driven by a compositional effect (e.g., all the immigrant students underperform vis-à-vis the U.S.-born students, mechanically increasing their score), in robustness analysis we repeat the same specification and standardize scores using only our sample of U.S.-born students. The results are substantially the same.

[14] In the Appendix, Table A3 we also clustered at the family level to take into account within family correlation; we also present the standard error when we cluster at the school level to take into account within school correlation across different cohorts. The results show slightly different standard errors, but our overall conclusions do not change.

observations is slightly lower). The math coefficient is now positive and significant, albeit very small, and the reading coefficient becomes positive, but statistically insignificant.

These measures of SES do not fully control for the selection of U.S.-born students who might select schools with small fractions of poor and immigrant students. We improve upon this specification by introducing a family fixed effect, ϕ_f, and compare across siblings.[15] The cumulative immigrant exposure coefficient becomes positive and statistically significant for both reading and mathematics. Indeed, the beta coefficient more than triples between the specifications in Columns 3 and 4 in both regressions. In Column 5, we present our most robust specification where we interact the family fixed effects with calendar year dummies, ϕ_{ft}, to control for life-cycle family trends in the same year and the results do not change substantially.[16] This novel result is driven by the introduction of family-fixed effects which take care of the selection problems due to sorting into schools based on immigrant concentration and SES.[17]

Our identification relies on the comparison between siblings in the same year, based on their "historic" exposure (including contemporary exposure): the variation may come from siblings going to different schools or different grades in the same school at a given point in time, but also from differences in past exposure. Thus, the sources of variation are: (1) siblings going to different grades in the same school(s) over time; (2) siblings going to different schools over time. The family-year fixed effects eliminate selection concerns when the source of variation is siblings going to different grades in the same schools (the regressions also include grade fixed effects). When the source of variation comes from siblings going to different schools over time, endogeneity concerns may arise: if the family selects different schools for different children, one may wonder if there is a kid-specific path over time that is not taken into account by our specification. We investigate this potential selection in Section 3.2.

[15] As we include family fixed effects, we remove the controls for ethnicity/race, lunch status, and mother's education. Miller, Shenhav and Grosz (forthcoming) show that the external validity of estimates obtained relying on within-family variation might be limited if the research design suffers from "selection into identification." We provide descriptive evidence that our results are not likely to suffer from this problem. We implement the observable-based reweighting procedure of Miller et al. (forthcoming) and we find that selection into identification is not a concern here (See Appendix B for details on the procedure used and the results).
[16] The time-varying controls could be endogenous. We rerun Table 4 excluding all the time varying controls (special education needs, free lunch eligibility and mother education) and the results stay the same (Table A4).
[17] Figlio and Ozek (2019) also use a sibling analysis as a robustness in their analysis of the contemporaneous correlation between an unexpected influx of Haitian refugees and the score of the incumbent students. They find a null effect. Their main result is different and not generalizable to immigrants in general as Haitian refugees' characteristics are different (most Haitian refugees returned to Haiti after a year, they were forced to leave their country of origin, and were poorer and eligible for government programs).

To illustrate how much variation is captured by the different fixed effects, in the Appendix, in Figure A1 we plot the distribution of the residuals for the cumulative immigrant exposure with four different models. In green, we plot the distribution of the demeaned exposure measure (Model 0), in red we plot the distribution of the residuals for the model including school-year and grade-year fixed effects (Model 1), in blue we plot the distribution of the residuals after partialling out school-year, grade-year, and family fixed effects (Model 2), and in yellow the residuals for the specification including school-year, grade-year, and family-year fixed effects (Model 3). While the family (and family-year) fixed effects capture a large portion of the variation, there is substantial variation left in the residuals to estimate a meaningful parameter. Figures A2 and A3 in the Appendix show the remaining variation of our outcomes of interest (math and reading scores respectively), after the inclusion of different sets of fixed effects. As for the residual variation in immigrant exposures, we still have enough variation left to estimate our parameter of interest.

To understand the economic magnitude of our effects, we compare our estimates to the relationship between maternal education and student outcomes. The standardized beta coefficient (0.012) of immigrant exposure in Column 5 is equal, for mathematics, to 8.5 percent of the difference in standardized test scores between students whose mother does not have a high school diploma and students whose mother has an high school diploma (the standardized beta coefficient in this case is 0.143).[18] The standardized beta coefficient of immigrant exposure on reading scores (0.006) is lower than math and corresponds to 4 percent of the difference in standardized test scores between students whose mother does not have a high school diploma and students whose mother has an high school diploma. Another way to calculate the economic significance is to compute the impact of moving from low cumulative exposure to high cumulative exposure. Moving from the 10th to the 90th percentile in the distribution of cumulative exposure (1 percent and 13 percent, respectively) would increase the score in mathematics and reading by 2.8 percent and 1.7 percent of a standard deviation, respectively.[19] We also study whether these magnitudes are different across grades by plotting the coefficients of the immigrant share interacted with grade in the baseline specification (Figure A4). We

---

[18] The excluded groups are mothers who are high school dropouts.

[19] Lavy and Schlosser (2011) study the effects of female classmates on students' academic achievements in Israel. They find that a 20-percentage-points increase in the proportion of female students translates into 4-5 percent of a standard deviation increase in test scores for both boys and girls in high-school. In our context, a standard deviation increase in cumulative exposure to foreign-born students roughly corresponds to a 5 percent increase in female share in Lavy and Schlosser (2011).

always find a positive and significant effect of similar magnitudes, except for a few grades for which the coefficient is not precisely estimated.

Taken together, these results suggest the presence of a strong selection of U.S.-born students into and out of schools potentially tampering the interpretation of regression results that do not control for sorting. To study whether this sorting is driven by specific sub-populations of U.S.-born students, in Tables 5 and 6, we split the sample by race/ethnicity and SES. In Tables 5A and 5B, we divide the sample into White and Black U.S.-born students and examine their performance in mathematics and readings.[20] The conditional correlation between immigrant exposure and the performance of White U.S.-born students is similar to Table 4: without the inclusion of any family control, it is negative and significant, but becomes positive with the inclusion of family fixed effects (significant for math and insignificant for reading). However, when we include the family-year fixed effects, the standardized beta coefficient is no longer significant. The results for Black students are quite different: the conditional correlation between immigrants' exposure and performance is stable and positive, independently of the controls included in the analysis. These results are complementary to the existing literature on "white flight" (e.g., Betts and Fairlie, 2003) because they suggest that White U.S.-born higher-performing-students are more likely to select into schools with a low fraction of minority and immigrant students. On the contrary, U.S.-born Black students do not select specifically into schools based on immigrant shares.

Since Black students are on average less affluent than White students, in Tables 6A and 6B, to further validate this interpretation, we separate higher and lower-SES U.S.-born students using FRPL eligibility. The results show that higher-SES students select into schools with a lower fraction of immigrants: the effect of immigrants is negative and significant when family controls are not included and becomes negligible and statistically insignificant, when family background is accounted for.[21] Conversely, the results for lower-SES students show that this group does not suffer from selection issues, similarly to Black students, and the coefficient is positive and significant in every specification.

Notably, Tables 5 and 6 show that the impact of immigrant exposure has differential effects on different subgroups. Compared with the overall sample, the effect of immigrant exposure is twice

[20] Black students are disproportionately poorer compared to White students in Florida public school system. Using Census data from 2010, we find that the average family income of Black students going to public school is only 61% of the average income of White students going to public schools. We only consider the subsamples of Black and White students because the sub-samples of Asian and Hispanic students are not large enough to estimate the coefficient of interest. The sub-sample of Hispanic students is significantly reduced by the restriction we impose on the language spoken at home.

[21] The reading coefficient become significant at the 10% level when we include family-year fixed effects.

as large for Black and FRPL-eligible students, while for White and FRPL ineligible students the effect is null and not significant.

3.2 Robustness Checks

In this section, we investigate whether our results are robust to alternative specifications and consider different ways to deal with potential selection of students into and out of schools.

3.2.1 Alternative Samples and Specifications

Several papers in the education literature have argued that the effects of time-varying inputs (schooling-related as well as child- and family-related) may decay over time rather than only be observed contemporaneously (Clotfelter et al., 2006; Clotfelter et al., 2007; Todd and Wolpin, 2003; Rothstein, 2010). Therefore, as a robustness for our measure of immigrant exposure, we consider a more general model using a geometric specification with different rates of decay, where λ represents the decay factor. For each student *i,* in school *s*, current grade *g,* and academic year *t*, the measure of cumulative exposure (weighted by distance in time from the current observation) is calculated using the following formula:

$$E_{isgt} = \frac{\sum_{g' \leq g} Immigrant\ Share_{isg't} \times e^{1-\left(\lambda(g-g')\right)}}{\sum_{g' \leq g} e^{1-\left(\lambda(g-g')\right)}} \qquad (3)$$

This specification permits a wide range of models, from our baseline model in which last year's exposure is just as influential as contemporaneous exposure (λ=0) to a model in which only contemporaneous exposure matters (λ increasing to infinity). The literature does not provide a direction on the specific size of λ: the previous literature has produced some estimates regarding decay in teachers' effect, but nothing specific regarding the effect of classmates. We re-calculate the measure of cumulative exposure by using different λ in equation (3) and re-estimate the model of Table 4, Column 5. Figure A5 plots the coefficients of cumulative exposures for different λ. We find highly consistent estimated effects of immigrant exposure regardless of the value of λ, suggesting that, in our specific application, the choice of λ does not drive our findings.

Our results are also robust to five additional specifications. The first pertains to the treatment of Puerto Rican-born students. As discussed previously, we re-run our specification, excluding Puerto Rican-born students from the immigrant groups. The results are reported in Appendix Table A5 and are consistent with our previous findings and the beta coefficients have similar magnitudes. The second robustness includes third grade scores as a control and recalculates the measure of immigrant exposure from third grade to the current grade. In this analysis, the sample size drops by roughly 60 percent. However, the interpretation of the results is remarkably similar (Table A6), given the slightly different specification which controls for initial conditions. In the third robustness, we re-run the specifications of Table 4, using a different definition of U.S.-born students, including students who do not speak English at home. The results are quantitatively similar to our main specification for math and readings (Tables A7). Fourth, it is possible that our estimated coefficient captures the presence of high achieving second generation students, if immigrants of various generations bunch together in the same schools (Card et al., 2000). To investigate this possibility, we add to our baseline regression a separate cumulative exposure measure for second generation immigrants, defined as students with mothers born in a different country. In our preferred specification, the main coefficient does not change substantially, while the second-generation immigrant exposure coefficient is insignificant (Table A8). Finally, to deal with children of military personnel who may be classified incorrectly as immigrants because they were born abroad while the parents were serving abroad, we exclude observations from the four Florida counties (Bay, Brevard, Clay, and Okaloosa) with large military concentrations. The results remain highly consistent regardless of whether we include or exclude these military-intensive counties (Table A9).

3.2.2 Additional Compositional and Selection Issues

While our most conservative estimate includes family-year fixed effects, which control for family lifecycle changes, one worry is that the results are mostly driven by the subset of siblings who go to different schools and by certain families whose children are very distant in years. To address this possibility, we run our baseline regressions for the sub-sample of siblings attending the same school. We first select families with only two children and then we divide this sample into those families whose children go into the same school in a given year and those who do not. The first sub-sample has siblings who are much closer in age, on average a difference of twenty versus 34 months. The results are presented in Table A10. Column 1 repeats our preferred specification with the sub-sample of all families with only two children, column 2 presents the results with children going to the same school, and column 3 shows the results with children going to different schools. If anything, the results seem to be stronger for math in the subsample of children going to the same school, while the coefficient is similar in all subsamples for reading but it is imprecisely estimated.

It is worth noting that our sibling comparison approach relies on the assumption that families make their school choice decisions independently of child-specific characteristics. By contrast, if parents were to send the highest achieving child to a school with fewer immigrants, the estimated coefficient on the share of immigrants would be downward biased. Alternatively, if parents have egalitarian preferences as in Becker and Tomes (1976) and believe that exposure to low-SES students and immigrants have a negative effect on their children performance, they may send the lower achieving child to a school with fewer immigrants. In this case, the estimated coefficient could be upward biased. Because school choice programs (e.g., open enrollment, charter schools) have become increasingly popular in Florida during the period of our study, this is a real possibility in our analysis.

To address the within family selection, we design an instrumental variable strategy. Families may select different schools for their children either by choosing a different school at the beginning of the academic cycle, or because, after choosing the same initial school, they select an alternative path for their children. We first address the latter case by accounting for possible family selections of different school paths for siblings who started in the same initial school (in possibly different years/grades). This sub-sample of students, roughly 67 percent of the sample, includes more stable families who do not move. Indeed, this sample is highly selected along academic achievement and various socio-economic characteristics. For the subset of siblings who go to the same initial school, the average math score is 0.192, the fraction FRPL-eligible is 45 percent, the fraction of White/Black students is 68 percent/22 percent. Maternal education is also higher for the students in this group: fewer students have mothers who dropped out of high school (15 percent), while more students have mothers who completed 4 years of college (24 percent).[22] Since Table 5 and 6 show that the positive effect of immigrants tends to affect U.S.-born students with lower SES, this selection issue works against finding an effect of immigrant exposure on U.S.-born students.

Using all the FLDOE data between 2002-2011, we construct for the entire population of students a transition matrix from school to school (grade by grade). Then, for each student in our IV

[22] In this IV estimation, we ignore the families who sends their kids to different initial schools which will be included in the second IV strategy. To study the motives behind the decision of having the second child in a different school, we analyze the sample of families with two children both attending elementary school (up to fifth grade) at the time in which the younger sibling enrolls in first grade. Among these families, 69 percent chose the same identical first school for both siblings in grade 1; 24 percent sent the two siblings to a different first initial school, but the first school of the younger sibling is the same as the current school of the older sibling. This latter statistic suggests that when the first initial school is different across siblings, it is generally due to the decision of the family to transfer all children to a new school, probably due to residential relocation, rather than due to a choice based on children's attitudes. The remaining 7 percent go to a first school which is different from the contemporaneous school of the older sibling.

sample, starting in each initial school, we use the school-to-school transition matrix to calculate the transition probabilities for each pair of consecutive grades. More formally, the transition matrix from grade *g* to grade *g+1,* is given by:

$$P(g+1|g) = \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} \dots\dots\dots \pi_{1N_s} \\ \pi_{21} & \pi_{22} & \pi_{23} \dots\dots\dots \pi_{2N_s} \\ . & . & . \qquad . \\ . & . & . \qquad . \\ \pi_{N_s1} & \pi_{N_s2} & \pi_{N_s3} \dots\dots\dots \pi_{N_sN_s} \end{bmatrix}$$

where π_{kj} is the probability that a student in school *k* at grade *g* ends up in school *j* at grade *g+1,* and N_s is the total number of schools in the sample.

We then multiply these transition probabilities with the fraction of immigrants observed in each potential school. Defining the set of different transition probabilities for the complete set of schools:

$$\left\{P(g+1|g)_{(N_s \times N_s)}\right\}_{g=0}^{11}$$

and the fraction of foreign students in a given school-grade-academic year:

$$\left\{\left\{W(g,t)_{(N_s \times 1)}\right\}_{g=0}^{12}\right\}_{t=2002}^{2011}$$

The predicted exposure at $(\tilde{g},\tilde{t})$ based on Markov chains for given (g_0,t_0) is given by:

$$Z(\tilde{g},\tilde{t})_{(N_S \times 1)} = E[W(\tilde{g},\tilde{t})|(g_0,t_0)] = \left(\prod_{g=g_0}^{g-1} P(g+1|g)\right)_{(N_S \times N_S)} W(\tilde{g},\tilde{t})_{(N_S \times 1)}$$

Overall, in our model, two siblings will have the same transition matrix but a different exposure to immigrants depending on the specific cohort they are in. Since our sample only includes families whose siblings started in the same school, we include grade time year and family by initial school fixed effect (family by year fixed effect would capture the full variation in immigrant exposure).

The binned scatter plots for the first stage, based on the unconditional model, and the model including family and initial school fixed effects are presented respectively in Figures 6A and 6B and provide evidence that the first stage is strong. The small difference between the actual exposure and the predicted exposure indicates that there is little within family selection in this sample. Table 7 further confirms the lack of differential selection within the family by reporting the full results of the IV together with the OLS and the reduced form for the same sample. The OLS coefficient of immigrant exposure is positive and significant. The instrumental variable coefficient is almost identical confirming that school choice is mostly done at the family level. We find equivalent results for reading.

This IV strategy does not address potential selection of families sending their children to different initial schools and excludes children if they are the third born or higher. While our analysis shows that most families who send their kids to different initial schools do so because the whole family has relocated, we design an alternative IV strategy, which also includes families sending their children to different initial schools. We use as an instrument the cumulative exposure the students would have had if she had gone to the same school of her oldest sibling. Our findings are consistent (Table A11 in the Appendix).[23]

4. Interpretation

In this section, we investigate potential explanations for our results, focusing on two main aspects that may correlate with the presence of immigrants: schools' responses and characteristics of the immigrants.

Institutional responses and schools' characteristics

The fraction of immigrant students can influence school resources, causing spillover effects for non-immigrant students. Indeed, at the federal level, schools receive additional funding, if immigrant students have educational needs as English language learners [ELLs] or special education students. At the school level, we cannot directly test the resource channel, but we can nevertheless rule it out because our specifications include school year fixed effects. At the classroom level, there are several potential mechanisms through which the resource channel may affect the results. Additional resources may change teacher characteristics and class size by cohort, potentially increasing student achievement (Krueger, 1999). We add to our specification cumulative average teachers' experience for each subject (math and reading).[24] In Table A12, column 1 and 4, we report the baseline

[23] To construct this alternative instrument, we attach to the younger sibling the school of the older sibling in the corresponding grade. Thus, we retain only the observations for which we observe siblings in the same grade. Because of this restriction, we are more likely to select siblings that are close in years. The coefficient of the IV is significantly larger than the OLS in this specification, probably because the 2SLS is estimating a local average treatment effect that puts additional weight to the subgroup of families that on average send their kids to the same school of the older sibling. These latter families are generally less interventionists, and their kids are likely the ones that benefit the most from the presence of immigrants. This result is consistent with our findings of a larger effect for low SES and minority students who, unlike White and affluent families, do not respond to the presence of immigrants by moving to different schools.

[24] Because students can have multiple math and reading classes, teacher experience is the weighted averages across teachers by subject calculated using as weights the fraction of time the student spends with each teacher. Note that we have fewer observations than in the main analysis for two reasons. First, teacher information is not available for every student in the enrollment file: classroom roster checks are conducted two times during

specification of Table 4, Column 5 (including family-year fixed effects). The results show that cumulative exposure to more experienced teachers is positively correlated with academic performance, but the inclusion of this control does not affect our main results. In Florida, similarly to many other states and school districts, class size is endogenous as prior achievement plays an important role in determining current year class size due to acceleration or remediation policies.[25] With this caveat in mind, in Column 2 and 5, we include cumulative average class size (respectively for math and reading): while the correlation between performance and class size is positive, our coefficient of interest remains unchanged. Finally, the results in Column 3 and 6, where we include both cumulative average teachers' experience and cumulative average class size, confirm that our observed effect is not driven by selective allocation of school resources in classes with a higher fraction of immigrants.

When there are more immigrant students attending the same school, they could be "segregated" in special classes, for example because these students take remedial English classes while U.S.-born students attend separate classes with potentially better targeted resources. If this is true, we should expect higher classroom segregation in schools with more immigrants and stronger outcomes when immigrant students are segregated in specific classrooms. To investigate if this is the case, we construct the following measure of segregation for each school, year, and grade:

$$Segregation_{syg} = \sum_{c \in syg} \left| \frac{FB_c}{FB_{syg}} - \frac{USB_c}{USB_{syg}} \right|$$

where FB_c is the number of foreign-born students in each classroom, FB_{syg} is the number of foreign-born students in the school, year, and grade, USB_c is the number of U.S.-born students in each classroom in the school, USB_{syg} is the number of U.S.-born students in the school. We first present in Figure A6 the correlation between the percentage of foreign-born students in the school, year, grade, and the segregation index. Differently from the hypothesis above, the larger the fraction of foreign-born students, the lower the amount of segregation. In Table A13, we present a regression analysis in which we explore what role segregation plays in our results. In column 1, we weight the

---

the school year (one in October and the other in February). Second, we do not include high school students because in high school, course selection is endogenous and depends on pre-qualification of students (for example, some students can take advanced-college level courses such as AP if they prequalify) and on whether students have selected a specific track (for example, International Baccalaureate, which is offered by some schools only).

[25] For example, during our sample period, middle school students in Florida were required to take remedial courses (in addition to the regular course) in ELA or math, which were required to be smaller in size, if they scored below the proficient level on the prior year test in that subject (Figlio and Ozek, forthcoming and Ozek, 2021).

cumulative exposure coefficient by segregation and re-estimate the column 5 model of Table 4 for both math and reading scores. The results show a beta coefficient slightly smaller for both math and reading but not significantly different from our baseline regression. Then, we compute the level of segregation in the contemporaneous school (columns 2 and 3) and in the first school the student was enrolled in (columns 4 and 5) and we split into the subsample of schools with above (columns 2 and 4) and below (columns 3 and 5) median segregation levels. Using these sub-samples, we find that the positive effect of immigrant exposure is not concentrated in the schools with higher segregation of immigrants. The standardized beta coefficient is always higher in the sub-samples of schools with lower segregation. This analysis indicates that immigrant segregation within schools is not responsible for our results.

Furthermore, the presence of immigrants may affect, positively or negatively, diversity in school. From a theoretical perspective, the effect of diversity is not obvious. Terenzini et al. (2001) finds evidence that the relation between the racial/ethnic composition of a classroom and students' learning gains is non-linear: medium levels of diversity are positively correlated with students' academic achievement, while low and high levels maybe associated with negative learning results. Empirically, it is an open question whether in our sample more immigrants lead to higher diversity. We create a measure of diversity, for the entire population of students at the school-grade-cohort level, as one minus the Herfindahl index by ethnicity/race (Alesina et al., 2003), and compute, for each child-observation, a cumulative exposure following equation (1). The correlation between the cumulative percentage of foreign-born students and the cumulative exposure to diversity is positive (25 percent) and significant at the 1 percent level, suggesting that U.S.-born students exposed to a large fraction of immigrants also experience more overall school diversity. Another measure of diversity pertains to the homogeneity of the immigrant population. Following the same procedure as above, we calculate two additional measures of diversity based on the population of immigrant students: one by race and the other by country of origin. Similarly, to the findings related to overall school body diversity, we find that when a U.S.-born student has a high cumulative exposure to immigrants, s/he is more likely to be exposed to immigrants with a diverse background (the correlation with immigrant diversity by ethnicity/race is 20 percent and by country of origin is 28 percent, both significant at the 1 percent level).

In the Appendix, Table A14, we investigate whether adding these three diversity measures into our baseline specification changes our main coefficient of interest. In all these specifications, our main explanatory variable is unaffected, suggesting that our measure of immigrant exposure is not a proxy for diversity in school. Overall, these results suggest that school policies, characteristics, and class composition are unlikely channels behind the positive correlation between the presence of immigrants and the performance of U.S.-born students.

Immigrants' characteristics, traits, and behavior

Immigrants belong disproportionally to racial minority groups, lower-SES families, and have limited English proficiency. Our immigrant exposure may, in principle, be a proxy for these characteristics. To study whether the presence of immigrants and the achievement of U.S.-born students is reflecting these socioeconomic characteristics of the immigrants, in Table 8, we saturate our model introducing a vector of cumulative exposures to racial minority groups, to students FRPL-eligible, with limited English proficiency, and receiving special education. These cumulative exposure variables are calculated as in equation (1) computed as leave-out-means. Even in this saturated specification, the coefficient on immigrant cumulative exposure remains statistically significant, with a larger magnitude (almost twice as large in reading and almost 50 percent larger in math), suggesting that it is the presence of immigrants and not their SES characteristics that indeed drives our findings. On the other hand, we find a negative correlation between academic achievement and exposure to students with low SES. The largest negative effect comes from being exposed to students receiving free or reduced lunch: the standardized beta coefficient of the exposure to students FRPL-eligible is -0.055 (-0.050) for mathematics (reading). Given this negative effect of low SES, it is natural to test whether the effect of immigrant exposure differs depending on the SES of the immigrants. In Columns 3 and 4 of Table 8, we augment our specification by creating two immigrant exposure variables, one measuring exposure to high SES immigrants (FRPL-ineligible) and the other measuring exposure to low SES immigrants (FRPL-eligible). Contrary to our previous results, it is the low SES immigrants that have the strongest effect on US born students.

A potential explanation for our main results is that the behavior of the immigrants has spillovers on their classmates either because they set examples through their behavior or performance or because they make the classroom environment more conducive to learning. To study whether the academic performance of immigrants drives our results, we need to address one major challenge. The reflection problem (Manski, 1993) prevents us from including the *absolute* performance of immigrant students in the regression because it may be affected by the performance of U.S.-born students. To address this problem, we substitute the absolute performance of immigrant students with a measure of the average immigrant academic performance by country of origin. This strategy relies on the assumption that the expected individual performance of a given immigrant is well proxied by the average performance of the immigrants coming from the same country of origin. Previous research suggests that the performance of immigrant students from the same country of origin is similar, independently from the country of destination and the school characteristics (Figlio et al., 2019), probably due to attitudes driven by values and beliefs in the country of origin. We use these measures of expected academic performance both in math and reading to weight our cumulative exposure to immigrants and add these immigrant performance indexes to our baseline regressions.[26] Table 9 shows the results for math and reading scores when we include these immigrant performance indexes into our analysis. In our preferred specification, with the inclusion of family-year fixed effects (column 1), this weighted index has a positive and significant coefficient, with a similar economic magnitude to the immigrant exposure's coefficient for math and double the magnitude for reading. The size of the immigrant exposure coefficient does not change compared to our baseline specification.

According to the literature, academic performance can also be affected by the level of disruption in the classroom. This effect could be driven by imitation or by an improved learning environment (Lazear, 2001; Carrell and Hoekstra, 2010; Carrell, Hoekstra, and Kuka, 2018). We measure disciplinary behavior using a dummy variable indicating whether the student was involved in a disciplinary incident during the school year (serious offense, often resulting in an in-school or out-of-school suspension). To study the potential impact of disciplinary behavior of immigrants, Table 10 repeats the same exercise as Table 9 by constructing an immigrant performance index based on the average disciplinary behavior of the immigrants by country of origin[27] (column 1). Exposure to better-behaved immigrants is associated with better academic outcomes. As before, this channel does not

[26] Our immigrant performance index is given by $\sum_c Immigrant\ Share_{cisg't} \times Performance_c$, where $Performance_c$ is the average math performance in the overall FLDOE data by country of origin, c, and $\sum_c Immigrant\ Share_{cisg't}$ is the sum of the share of immigrants in school s, grade g, at time t that each U.S.-born student i observes (the sum of the shares of immigrants is equal to one). The distribution of the country-of-origin performances (plotted in Figure A7, Panel A and B in the Appendix) confirm major differences among countries of origin in math and reading. Immigrant exposure is negatively correlated (-0.22) with the immigrant performance index (in areas where there are more immigrants, the average academic achievement of the immigrant is lower). This measure addresses the reflection problem, as the performance of each immigrant group is calculated at the population level. Nonetheless, because the national origin mix of immigrants in a school is non-random, there is some potential endogeneity in this measure.

[27] The distribution of the country-of-origin disciplinary behavior (Figure A7, Panel C in the Appendix) confirm major differences among countries of origin. Immigrant exposure is positively correlated (0.16) with the immigrant performance index based on disciplinary incidents (in areas where there are more immigrants, the average behavior of the immigrants is better).

affect the direct impact of immigrant exposure: the beta coefficient of this variable remains similar to the baseline specification.[28]

When we split the sample by SES, we find some interesting results in comparison to Tables 5 and 6. The immigrant performance indexes based on math and reading scores have very similar effects in all subsamples, suggesting that the absolute performance of the immigrants has a consistent positive effect on all U.S.-born students, unlike our main variable (immigrant exposure) which behaves exactly as in Tables 5 and 6: positive and significant for FRPL-eligible and Black students, null and insignificant for White and FRPL-ineligible students. When we use the immigrant performance index based on disciplinary incidents, the results are similar.

Overall, these results suggest that the presence of immigrants with higher academic performance correlates with better scores for all U.S.-born students. However, even after controlling for the absolute performance of immigrant students, a higher exposure to immigrants is still associated with higher achievement of U.S.-born students, concentrated among less affluent and Black U.S.-born students.

To explain the differential impact of immigrants on high and low socio-economic status U.S.-born students, we analyze the *relative* academic performance and behavior of immigrant students vis-à-vis their schoolmates belonging to different subgroups. In Table 11 we compare math, reading scores, and disciplinary behavior between U.S.-born students and corresponding immigrant students going to school with them for different SES sub-samples. We find that the immigrants (average math score -0.137) substantially outperform low SES U.S.-born students (average math score -0.303). We also find that immigrants who attend schools with low SES U.S.-born students have fewer disciplinary incidents (0.131 versus 0.247). The gap between immigrants and U.S.-born Black students is even larger. Reading scores show very similar trends. The comparison in performance and behavior between immigrants and U.S.-born students is completely different for the subset of affluent and White students. Immigrants going to school with high SES U.S.-born students have an average math score of 0.170 and underperform their U.S.-born schoolmates (0.475). Similarly, immigrants going to school with White U.S.-born students have a score in math of 0.093, while their White schoolmates have corresponding score of 0.305. The results on differences in disruptive behavior and reading scores are consistent with math scores.

[28] These results are unchanged when we saturate our model introducing a vector of cumulative exposures to racial minority groups, to students FRPL-eligible, with limited English proficiency, and receiving special education (Tables A15 and A16).

Overall, this descriptive evidence, combined with the results of Tables 5 and 6, suggests that the performance of minority and low SES U.S.-born students improves when they face a larger fraction of immigrant students that, in relative terms, perform and behave better than them. However, the performance of affluent U.S.-born students is not affected negatively by a large fraction of lower-performing immigrants. These results suggest that relative performance does not provide a clear channel for our results.

Hsin and Xie (2014) and Figlio et al. (2019), in different contexts, find that the behavior of immigrants can be explained by cultures that emphasize deferred gratification and self-control: these unobservable immigrant traits may help explaining our results. Figlio et al (2019) show that immigrant students with high long-term orientation succeed more in school than other immigrants. As a reference US students have a long-term orientation that falls in the bottom quartile of the world distribution. To test whether LTO's immigrant attitudes have positive spillovers on US born students, in Table 12, we split the exposure to immigrants distinguishing between immigrants with LTO above and below the US and added these measures to our specification.[29] We find that, independently of socio-economic status, the presence of immigrants with LTO higher than the US born, always benefit their US classmates. These results are suggestive that immigrants' attitudes, beyond their academic performance for which we control in the regressions, may play a role in explaining our results.

5. Conclusions

We study the effect of exposure to immigrants on educational outcomes of U.S.-born students using a large panel combining population-level administrative data from the Florida Department of Education Data Warehouse and birth records from the Florida Department of Vital Statistics. Our data allow us to use a novel identification strategy to deal with both the endogenous selection of immigrant students into schools and the endogenous response of incumbent families by comparing the test scores in math and reading of siblings who experience different school-cohort-specific immigrant concentrations, holding the heterogeneity of the families' lifecycles fixed.

Our main result points to a strong selection of U.S.-born students into and out of schools, potentially tampering the interpretation of regression analysis that do not control for this sorting mechanism. This selection problem is concentrated among White U.S.-born and higher-SES students consistently with the "white flight" literature.

[29] Since we do not have LTO data for Cuba, Jamaica, and Haiti, our specification also includes a residual exposure component for these three countries without LTO.

Our identification strategy provides new results about the effects of immigrants on the educational outcomes of U.S.-born students: once selection is accounted for with family fixed effects, the correlation between cumulative immigrant exposure and academic achievement of U.S.-born students is positive and significant. Moving from the 10^{th} to the 90^{th} percentile in the distribution of cumulative exposure (1 percent and 13 percent, respectively) increases the score in mathematics and reading by 2.8 percent and 1.7 percent of a standard deviation, respectively. The effect is double in size for disadvantaged students (Black and FRPL-eligible students). For affluent students, the effect is very small.

Even after controlling for absolute performance of immigrant students, a higher fraction of immigrants is still associated with higher achievement of U.S.-born students. We find that the presence of immigrant students benefits disadvantaged U.S.-born students, who, in relative terms, perform worse than their immigrant classmates. However, relative performance cannot fully explain our results: when immigrant students perform worse than their classmates, they do not have negative spillovers on them. To reconcile this latter finding, we investigate whether difficult-to-measure attributes of immigrants such as hard work and resilience play a role in our results. We isolate the importance of Long-Term Orientation, the disposition to sacrifice the present for the future, by dividing immigrant exposure into two groups, those with LTO above/below the US. We find that, controlling for school performance of the immigrants, our main effect is driven by high LTO immigrants. Overall, our evidence suggests that horizontal transmission of attitudes and beliefs may play an important role in explaining educational performance.

References

Alba, Richard, and John Logan (1993). "Minority Proximity to Whites in Suburbs: An Individual Level Analysis of Segregation." *American Journal of Sociology* 98(6): 1388-1427.

Alba, Richard, John Logan, Brian Stults, Gilbert Marzan, and Wengquan Zhang (1999). "Immigrant Groups in the Suburbs: A Reexamination of Suburbanization and Spatial Assimilation." *American Sociological Review* 64:3, 446-460.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003) "Fractionalization." *Journal of Economic Growth* 8: 155-94.

Ballatore R., M. Fort, and A. Ichino (2018), "The Tower of Babel in the Classroom Immigrants and Natives in Italian Schools," *Journal of Labor Economics.*

Becker, G., and N. Tomes (1976), "Child Endowments and the Quantity and Quality of Children," *Journal of Political Economy* 84: 143-162.

Betts, J., and R. Fairlie (2003) "Does Immigration Induce 'Native Flight' from Public Schools into Private Schools?" *Journal of Public Economics* 87: 987–1012.

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: 733-766

Brunello, G., and L. Rocco (2013), "The Effect of Immigration on the School Performance of Natives: Cross Country Evidence Using PISA Test Scores," *Economics of Education Review* 32: 234-246.

Card, D. (2001), "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics* 59: 22-64.

Card, D., DiNardo J., and E. Estes (2000), "The More Things Change: Immigrants and the Children of Immigrants in the 1940s, the 1970s, and the 1990s," in George Borjas (editor) *Issues in the Economics of Immigration*, University of Chicago Press, Chicago.

Carrell, S., and M. Hoekstra (2010), "Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids", *American Economic Journal: Applied Economics* 2: 211-228.

Carrell, S., M. Hoekstra, and E. Kuka (2018), "The Long-run Effects of Disruptive Peers", *American Economic Review* 108: 3377-3415.

Cascio, E. U., and E. G. Lewis (2012), "Cracks in the Melting Pot: Immigration, School Choice, and Segregation." *American Economic Journal: Economic Policy*, 4 (3): 91-117.

Clotfelter, C., H. Ladd, and J. Vigdor (2006), "Teacher-Student Matching and the Assessment of Teacher Effectiveness," *Journal of Human Resources* 41: 778–820.

Clotfelter, C., H. Ladd, and J. Vigdor (2007), "How and Why Do Teacher Credentials Matter for Student Achievement?," Working paper 12828, National Bureau of Economic Research.

Fairlie R. and A. Resch (2002), "Is there "White Flight" into Private Schools? Evidence from the National Educational Longitudinal Survey," *Review of Economics and Statistics* 84: 21–33.

Figlio, D., J. Guryan, K. Karbownik, and J. Roth (2014) "The Effects of Poor Neonatal Health on Children's Cognitive Development," *American Economic Review* 104: 3921-3955.

Figlio, D., P. Giuliano, U. Özek, and P. Sapienza (2019), "Long-Term Orientation and Educational Performance," *American Economic Journal: Economic Policy* 11: 272-309.

Figlio, D. and U. Özek (2019), "Unwelcome Guests? The Effects of Refugees on the Educational Outcomes of Incumbent Students," *Journal of Labor Economics*, 37 (4): 1061-1096.

Figlio, D. and U. Özek (forthcoming), "The Unintended Consequences of Test-Based Remediation," *American Economic Journal: Applied Economics*.

Fix, M., and W. Zimmerman (1993) *Educating Immigrant Children: Chapter I in the Changing City*, Washington: Urban Institute Press.

Freeman, Lance (2002). "Does Spatial Assimilation Work for Black Immigrants in the U.S.?" *Urban Studies* 39(11): 1983-2003.

Friedman, Samantha, and Emily Rosenbaum (2004). "Nativity Status and Racial/Ethnic Differences in Access to Quality Housing: Does Home Ownership Bring Greater Parity?" *Housing Policy Debate* 15(4): 865-901.

Geay, C., McNally, S. and Telhaj, S. (2013), "Non-native Speakers of English in the Classroom: What Are the Effects on Pupil Performance?" *Economic Journal* 123: F281-F307.

Gould, E., Lavy, V. and Daniele Paserman, M. (2009), "Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence," *Economic Journal* 119: 1243-1269.

Hsin, A. and Y. Xie (2014), "Explaining Asian-Americans' Academic Advantage Over Whites," *PNAS*, 111 (23), 8416-8421.

Hunt, J. (2016), "The Impact of Immigration on the Educational Attainment of Natives," *Journal of Human Resources* 52: 1060-1118.

Jensen, P., and A. Rasmussen (2011), "The Effect of Immigrant Concentration in Schools on Native and Immigrant Children's Reading and Math Skills," *Economics of Education Review* 30: 1503-1515.

Krueger, A. (1999), "Experimental Estimates of Education Production Functions", *The Quarterly Journal of Economics*, 114 (2), 497-532.

Lavy, V. and A. Schlosser (2011), "Mechanisms and Impacts of Gender Peer Effects at School." *American Economic Journal: Applied Economics* 3: 1–33.

Lazear, E. (2001), "Educational Production," *Quarterly Journal of Economics* 116: 777-803.

Logan, John, and Richard Alba (1993). "Locational Returns to Human Capital: Minority Access to Suburban Community Resources." *Demography* 30(2): 243-268.

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

McHenry, P. (2015) "Immigration and the Human Capital of Natives." *Journal of Human Resources* 50: 34–71.

Miller, D., N. Shenhav and M. Grosz (forthcoming), "Selection into Identification in Fixed Effects Models with Application to Head Start," *Journal of Human Resources.*

Morales, N. C. (2020), Do Refugee Students Affect the Academic Achievement of Peers? Evidence from a Large Urban School District. Available at SSRN: https://ssrn.com/abstract=3731794

Neymotin, F. (2009), "Immigration and its Effect on the College-going Outcomes of Natives," *Economics of Education Review* 28: 538–550.

Nunn, R., J. O'Donnell, and J. Shambaugh (2018) "A Dozen Facts about Immigration," The Hamilton Project.

Ohinata, A. and van Ours, J. C. (2013), "How Immigrant Children Affect the Academic Achievement of Native Dutch Children," *Economic Journal* 123: 308-331.

Özek U. (2021), "The Effects of Middle School Remediation on Postsecondary Success: Regression Discontinuity Evidence from Florida," *Journal of Public Economics*, 2021, 203, 104518.

Ozek U. (forthcoming), Examining the Educational Spillover Effects of Severe Natural Disasters: The Case of Hurricane Maria, *The Journal of Human Resources.*

Rosenbaum, Emily, Samantha Friedman, Michael Schill, and Hielke Buddelmeyer (1999). "Nativity Differences in Neighborhood Quality among New York City Households, 1996." *Housing Policy Debate* 10(3): 625-658.

Rosenbaum, Emily, and Samantha Friedman (2001). "Differences in the Locational Attainment of Immigrant and Native-Born Households with Children in New York Cit." *Demography* 38(3): 337-348.

Rosenbaum, Emily, and Samantha Friedman (2004). "Generational Patterns of Home Ownership and Housing Quality among Racial/Ethnic Groups in New York City." *International Migration Review* 38(4): 1492-1533.

Rothstein, J. (2010), "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement," *Quarterly Journal of Economics* 125: 175-214.

Schill, Michael, Samantha Friedman, and Emily Rosenbaum (1998). "The Housing Conditions of Immigrants in New York City." *Journal of Housing Research* 9(2): 201-235.

Schneeweis N. (2015) "Immigrant Concentration in Schools: Consequences for Native and Migrant Students," *Labour Economics* 35: 63–76.

Schwartz, A. and Stiefel, L. (2011). Immigrants and Inequality in Public Schools. In *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances* (pp. 419-441). Russell Sage Foundation.

Terenzini, P. T., Cabrera, A. F., Colbeck C. L., Bjorklund S. A., and Parente J. M. (2001) Racial and Ethnic Diversity in the Classroom, *The Journal of Higher Education*, 72:5, 509-531.

Todd, P., and K. Wolpin (2003), "On the Specification and Estimation of the Production Function for Cognitive Achievement," *Economic Journal* 113: F3– F33.

Tonello M. (2016), "Peer Effects of Non-native Students on Natives' Educational Outcomes: Mechanisms and Evidence," *Empirical Economics* 51: 383-414.

Van der Werf, C. (2021), "The Impact of Refugees on Native Students' Academic Achievement." Louisiana State University, mimeo.

Figures

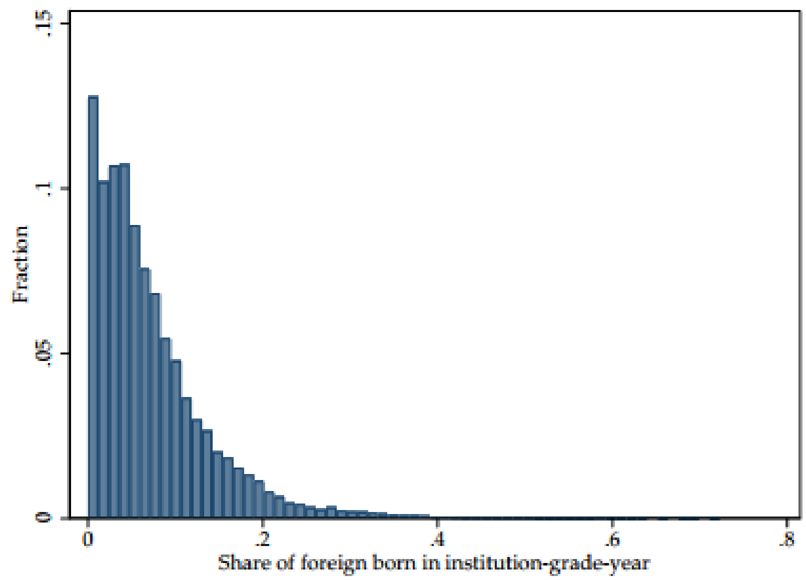

Figure 1: Distribution of foreign-born students as a share (a measure in the interval [0,1]) of the total number of students across institution-grade-year cells. The y-axis refers to the fraction of observations corresponding to values of the x-axis. The reference sample of U.S.-born students is an unbalanced longitudinal sample of U.S.-born students observed in grades from 3rd to 10th, who speak English at home and have at least one sibling, using observations in academic years in which at least two students are observed for each family.

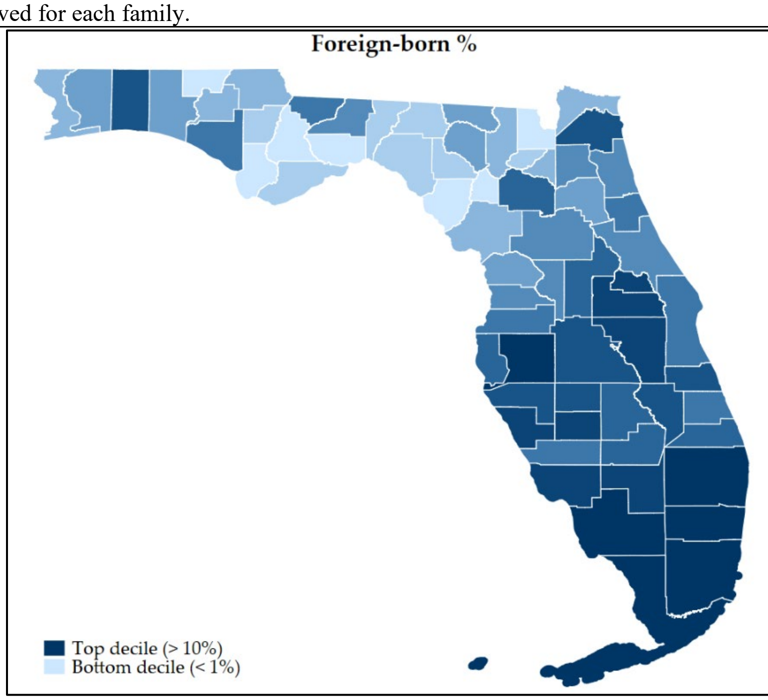

Figure 2: Concentration of foreign-born students across school districts in Florida. For each district we computed the percentage of foreign-born students over the total population of students across all years in the sample. The distribution across districts has been split in deciles and each gradation of blue corresponds to a decile in the distribution. Lighter blue indicates a lower percentage of foreign-born students, while darker blue indicates a higher concentration. The reference sample of U.S.-born students is an unbalanced longitudinal sample of U.S.-born students observed in grades from 3rd to 10th, who speak English at home and have at least one sibling, using observations in academic years in which at least two students are observed for each family.

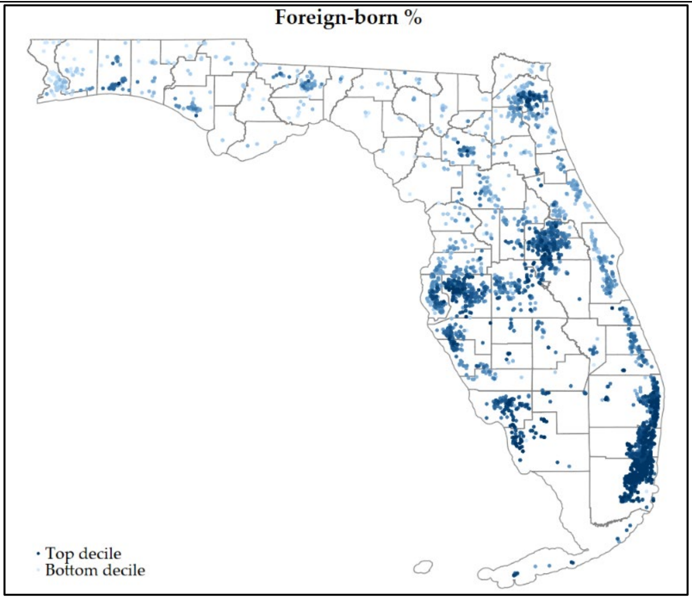

Figure 3A: Each dot in the map corresponds to one educational institution (a school). For each school we computed the percentage of foreign-born students over the total population of students across all years in the sample. The distribution across schools has been split in deciles and each gradation of blue corresponds to a decile in the distribution. Lighter blue indicates a lower percentage of foreign-born students, while darker blue indicates a higher concentration. The reference sample of U.S.-born students is an unbalanced longitudinal sample of U.S.-born students observed in grades from 3rd to 10th, who speak English at home and have at least one sibling, using observations in academic years in which at least two students are observed for each family.

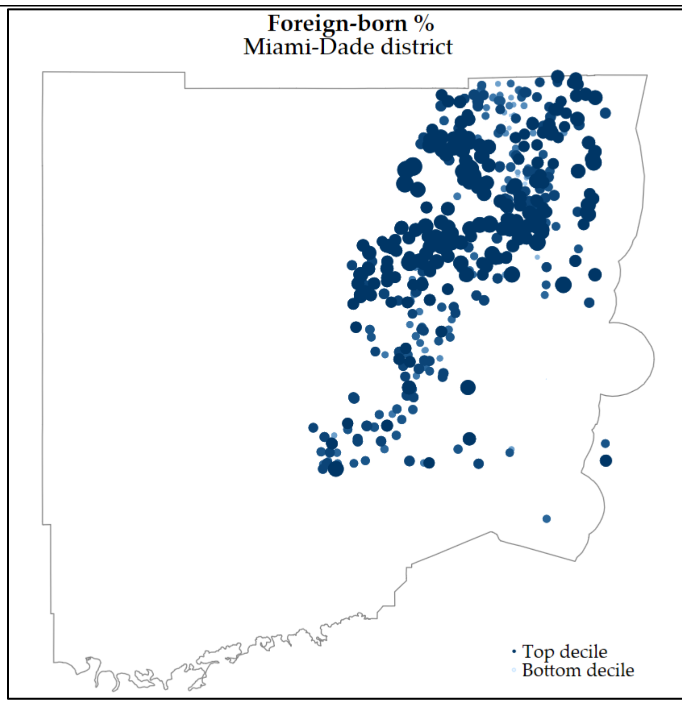

Figure 3B: Each dot corresponds to an educational institution in the Miami-Dade school district. The meaning of the color is the same as in Figure 3A: lighter colors correspond to lower deciles in the distribution of foreign-born students' concentration in the whole state of Florida. The size of the dots corresponds to the size of the student body. The reference sample of U.S.-born students is an unbalanced longitudinal sample of U.S.-born students observed in grades from 3rd to 10th, who speak English at home and have at least one sibling, using observations in academic years in which at least two students are observed for each family.

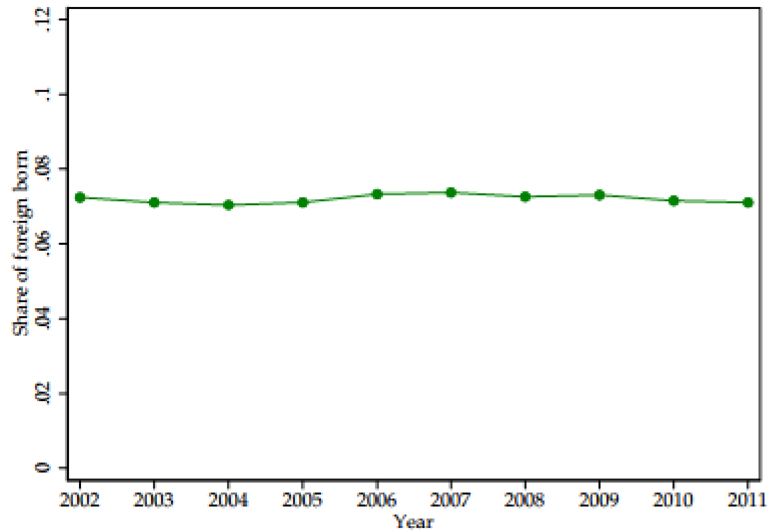

Figure 4: Using observations across the entire time span available in the data (2002-2011), we compute the average share of foreign-born classmates for U.S.-born English-speaking students, for each year from 2002 to 2011. The reference sample of U.S.-born students is an unbalanced longitudinal sample of U.S.-born students observed in grades from 3rd to 10th, who speak English at home and have at least one sibling, using observations in academic years in which at least two students are observed for each family.

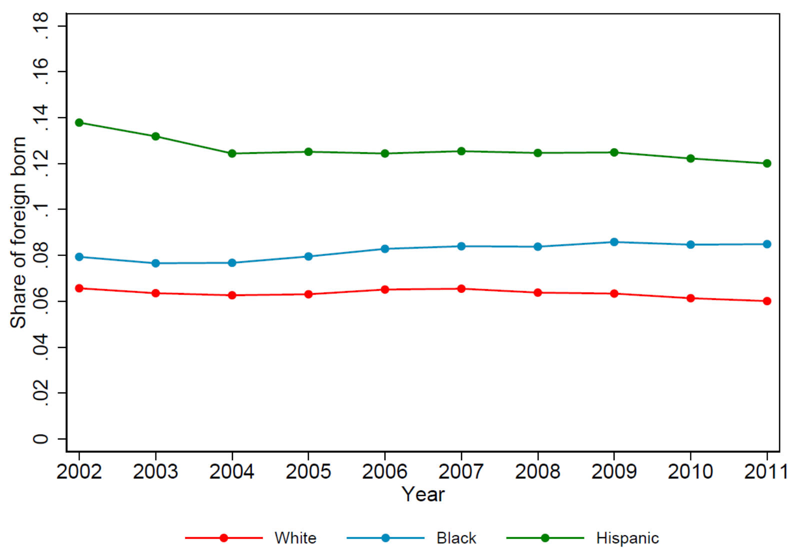

Figure 5: Using observations across the entire time span available in the data (2002-2011), we compute the average share of foreign-born classmates for the three major racial/ethnic groups of U.S.-born English-speaking students, for each year from 2002 to 2011. The red line shows average exposures to foreign-born students for White U.S.-born students, the blue line shows an analogous figure for Black U.S.-born students, and the green line shows an analogous figure for Hispanic U.S.-born students. The reference sample of U.S.-born students is an unbalanced longitudinal sample of U.S.-born students observed in grades from 3rd to 10th, who speak English at home and have at least one sibling, using observations in academic years in which at least two students are observed for each family.

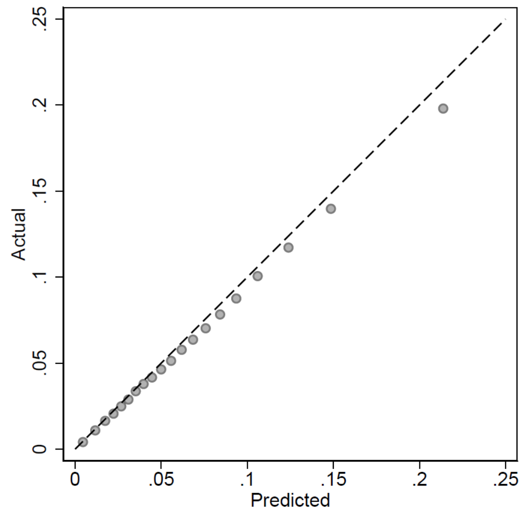

Figure 6A: This figure is a binned scatter plot that shows the raw correlation between the predicted cumulative exposure to foreign-born students and the actual cumulative exposure. Please refer to the text for details on the construction of the predicted cumulative exposure. The dashed line represents the 45-degree locus, along which the two variables are identical. The reference sample of U.S.-born students is an unbalanced longitudinal sample of U.S.-born students observed in grades from 3rd to 10th, who speak English at home. The sample is further restricted to students from families where all siblings attended the same initial school (i.e. the first school a student is observed in).

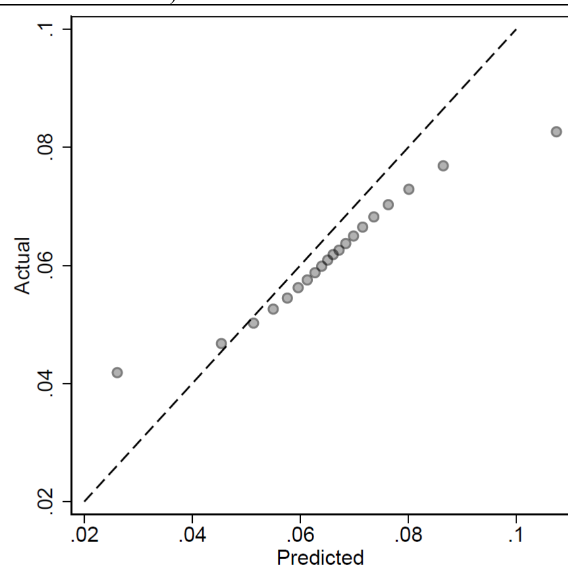

Figure 6B: This figure is a binned scatter plot that shows the correlation between the predicted cumulative exposure to foreign-born students and the actual cumulative exposure, conditional on family by initial school fixed effects. Please refer to the text for details on the construction of the predicted cumulative exposure. The dashed line represents the 45-degree locus, along which the two variables are identical. The reference sample of U.S.-born students is an unbalanced longitudinal sample of U.S.-born students observed in grades from 3rd to 10th, who speak English at home. The sample is further restricted to students from families where all siblings attended the same initial school (i.e. the first school a student is observed in).

Tables

Variable	Obs	Mean	Std. Dev.
Outcomes:			
Math Standardized Score	1,347,286	0.050	0.993
Reading Standardized Score	1,450,138	0.034	0.992
Incidents (ever involved in)	1,450,138	0.169	0.375
Explanatory variables of interest:			
Foreign-born Exposure	1,347,286	0.060	0.052
Foreign-born Exposure (Predicted)	821,892	0.066	0.052
Immigrant performance index (Math score)	1,271,246	-0.037	0.280
Immigrant performance index (Behavior)	1,271,246	0.143	0.032
Immigrant performance index (Reading score)	1,371,517	-0.150	0.267
Exposure to migrants with LTO above US	1,271,246	0.017	0.016
Exposure to migrants with LTO below US	1,271,246	0.026	0.030
Exposure to migrants with low SES	1,347,286	0.037	0.040
Exposure to migrants with high SES	1,347,286	0.023	0.026
Exposure to Black students	1,347,286	0.246	0.236
Exposure to White students	1,347,286	0.510	0.259
Exposure to Asian students	1,347,286	0.022	0.021
Exposure to Free-Lunch students	1,347,286	0.532	0.234
Exposure to Limited English Proficiency students	1,347,286	0.069	0.080
Exposure to Special Education Needs students	1,347,286	0.148	0.050
Individual or family characteristics:			
Female (Indicator)	1,347,286	0.498	0.500
Age in Months	1,347,286	135.5	23.2
Special Education (Indicator)	1,347,286	0.147	0.354
Birth Order	1,347,286	2.199	1.170
White (Indicator)	1,347,286	0.603	0.489
Black (Indicator)	1,347,286	0.297	0.457
Hispanic (Indicator)	1,347,286	0.052	0.223
Asian (Indicator)	1,347,286	0.007	0.082
Other (Indicator)	1,347,286	0.042	0.200
Free/Reduced-Price Lunch (Indicator)	1,347,286	0.546	0.498
Limited English Proficiency (Indicator)	1,347,286	0.002	0.043
Mother High School DO (Indicator)	1,344,541	0.200	0.400
Mother High School Graduate (Indicator)	1,344,541	0.367	0.482
Mother Some College (Indicator)	1,344,541	0.239	0.426
Mother 4-year College or more (Indicator)	1,344,541	0.194	0.396

Table 1: Summary statistics. Cumulative exposure to foreign-born students (foreign-born exposure) is computed as the average share of foreign-born students across previous school-specific cohorts including the current grade. All statistics are computed on an unbalanced longitudinal sample of U.S.-born students observed in grades from 3^{rd} to 10^{th}, who speak English at home and have at least one sibling, using observations in academic years in which at least two students are observed for each family.

		Overall		White Majority		Hispanic Majority		Black Majority	
1	Cuba	16%	Mexico	13%	Cuba	46%	Haiti	41%
2	Mexico	10%	Puerto Rico	7%	Colombia	9%	Jamaica	13%
3	Haiti	10%	Colombia	7%	Mexico	7%	Mexico	6%
4	Colombia	8%	Germany	5%	Venezuela	6%	Puerto Rico	4%
5	Puerto Rico	6%	Cuba	4%	Puerto Rico	4%	Cuba	3%
6	Venezuela	5%	Canada	4%	Honduras	3%	Honduras	3%
7	Jamaica	3%	Haiti	3%	Dominican Republic	3%	Dominican Republic	2%
8	Peru	3%	Venezuela	3%	Argentina	3%	The Bahamas	2%
9	Argentina	2%	Brazil	3%	Peru	3%	Colombia	2%
10	Honduras	2%	China	3%	Nicaragua	3%	Japan	1%
Top-10 Cumulative		65%		50%		85%		78%

Table 2: Top 10 countries of origin of immigrants in Florida facing our sample of U.S.-born students. White/Hispanic/Black majority indicates that only school-specific cohorts with more than 50% U.S.-born of that specific race/ethnicity are selected. The reference sample of U.S.-born students is an unbalanced longitudinal sample of U.S.-born students observed in grades from 3rd to 10th, who speak English at home and have at least one sibling, using observations in academic years in which at least two students are observed for each family. The cumulative percentages may not add up to the column total due to within-cell rounding.

		Overall		White Majority		Hispanic Majority		Black Majority
1	Hispanic	62%	Hispanic	46%	Hispanic	92%	Black	63%
2	Black	17%	White	29%	Black	3%	Hispanic	28%
3	White	13%	Asian	13%	White	3%	Asian	5%
Top-3 Cumulative		91%		88%		98%		95%

Table 3: Top racial/ethnic groups of immigrants in Florida facing our sample of U.S. born students. White/Hispanic/Black majority indicates that only school-specific cohorts with more than 50% U.S.-born of that specific race/ethnicity are selected. All statistics are computed on an unbalanced longitudinal sample of U.S.-born students observed in grades from 3rd to 10th, who speak English at home and have at least one sibling, using observations in academic years in which at least two students are observed for each family. The cumulative percentages may not add up to the column total due to within-cell rounding.

	Standardized scores (3rd-10th grade)				
	(1)	(2)	(3)	(4)	(5)
	Panel A: Math standardized score				
Foreign-born Exposure	-0.123**	0.019	0.077*	0.293***	0.229***
	(0.053)	(0.042)	(0.040)	(0.054)	(0.074)
	[-0.006]	[0.001]	[0.004]	[0.015]	[0.012]
Observations	1,347,286	1,347,286	1,344,541	1,347,286	1,347,286
R-squared	0.302	0.359	0.379	0.682	0.769
Mean LHS	0.0504	0.0504	0.0510	0.0504	0.0504
SD LHS	0.993	0.993	0.993	0.993	0.993
	Panel B: Reading standardized score				
Foreign-born Exposure	-0.194***	-0.026	0.040	0.176***	0.110*
	(0.049)	(0.039)	(0.037)	(0.048)	(0.064)
	[-0.010]	[-0.001]	[0.002]	[0.009]	[0.006]
Observations	1,450,138	1,450,138	1,447,278	1,450,138	1,450,138
R-squared	0.303	0.356	0.377	0.667	0.752
Mean LHS	0.0340	0.0340	0.0345	0.0340	0.0340
SD LHS	0.992	0.992	0.992	0.992	0.992
Individual Controls	X	X	X	X	X
School x Year FE	X	X	X	X	X
Grade x Year FE	X	X	X	X	X
Race FE		X	X		
Lunch Status		X	X		
Mother's Education FE			X		
Family FE				X	
Family x Year FE					X

Table 4: This table shows the estimates of a linear regression of test scores in mathematics (Panel A) and reading (Panel B) standardized by year and grade on the cumulative exposure to foreign-born students, computed as the average share of foreign-born students across previous school-specific cohorts including the current grade, and several controls. All regressions are run on an unbalanced longitudinal sample of U.S.-born students observed in grades from 3rd to 10th, who speak English at home and have at least one sibling, using observations in academic years in which at least two students are observed for each family. Individual controls include: gender, age in months, special education, and birth order fixed effects. Lunch status is a dummy variable equal to 1 if the student is eligible for free or reduced-price lunch. Mother's education fixed effects are three dummy variables equal to 1 if the mother of the student has a high school diploma, some college, or a four-year college or more, respectively. Robust standard errors in parenthesis clustered by school-cohort. Beta standardized coefficients in squared parenthesis below standard errors. *** p<0.01, ** p<0.05, * p<0.1.

	Standardized scores (3rd-10th grade)				
	Sample restriction: Race = 'White'				
	(1)	(2)	(3)	(4)	(5)
	Panel A: Math standardized score				
Foreign-born Exposure	-0.610***	-0.395***	-0.261***	0.213***	0.128
	(0.064)	(0.061)	(0.058)	(0.075)	(0.107)
	[-0.031]	[-0.020]	[-0.014]	[0.011]	[0.007]
Observations	811,790	811,790	810,559	811,790	811,790
R-squared	0.263	0.284	0.312	0.671	0.764
Mean LHS	0.305	0.305	0.305	0.305	0.305
SD LHS	0.911	0.911	0.911	0.911	0.911
	Panel B: Reading standardized score				
Foreign-born Exposure	-0.759***	-0.528***	-0.378***	0.044	-0.009
	(0.062)	(0.059)	(0.056)	(0.073)	(0.099)
	[-0.039]	[-0.027]	[-0.019]	[0.002]	[0.000]
Observations	873,281	873,281	872,002	873,281	873,281
R-squared	0.247	0.266	0.294	0.643	0.738
Mean LHS	0.288	0.288	0.288	0.288	0.288
SD LHS	0.933	0.933	0.933	0.933	0.933
Individual Controls	X	X	X	X	X
School x Year FE	X	X	X	X	X
Grade x Year FE	X	X	X	X	X
Race FE		X	X		
Lunch Status		X	X		
Mother's Education FE			X		
Family FE				X	
Family x Year FE					X

Table 5A: This table shows the estimates of a linear regression of test scores in mathematics (Panel A) and reading (Panel B) standardized by year and grade on the cumulative exposure to foreign-born students, computed as the average share of foreign-born students across previous school-specific cohorts including the current grade, and several controls. All regressions are run on an unbalanced longitudinal sample of U.S.-born White students observed in grades from 3rd to 10th, who speak English at home and have at least one sibling, using observations in academic years in which at least two students are observed for each family. Individual controls include: gender, age in months, special education, and birth order fixed effects. Lunch status is a dummy variable equal to 1 if the student is eligible for free or reduced-price lunch. Mother's education fixed effects are three dummy variables equal to 1 if the mother of the student has a high school diploma, some college, or a four-year college or more, respectively. Robust standard errors in parenthesis clustered by school-cohort. Beta standardized coefficients in squared parenthesis below standard errors. *** p<0.01, ** p<0.05, * p<0.1.

	Standardized scores (3rd-10th grade)				
	Sample restriction: Race = 'Black'				
	(1)	(2)	(3)	(4)	(5)
	Panel A: Math standardized score				
Foreign-born Exposure	0.517***	0.500***	0.481***	0.450***	0.402***
	(0.067)	(0.066)	(0.065)	(0.097)	(0.137)
	[0.028]	[0.027]	[0.026]	[0.025]	[0.022]
Observations	399,585	399,585	398,268	399,585	399,585
R-squared	0.266	0.273	0.283	0.593	0.716
Mean LHS	-0.495	-0.495	-0.495	-0.495	-0.495
SD LHS	0.951	0.951	0.951	0.951	0.951
	Panel B: Reading standardized score				
Foreign-born Exposure	0.563***	0.551***	0.533***	0.371***	0.286***
	(0.059)	(0.058)	(0.058)	(0.082)	(0.110)
	[0.033]	[0.032]	[0.031]	[0.022]	[0.017]
Observations	430,974	430,974	429,597	430,974	430,974
R-squared	0.286	0.296	0.307	0.593	0.707
Mean LHS	-0.511	-0.511	-0.511	-0.511	-0.511
SD LHS	0.904	0.904	0.904	0.904	0.904
Individual Controls	X	X	X	X	X
School x Year FE	X	X	X	X	X
Grade x Year FE	X	X	X	X	X
Race FE		X	X		
Lunch Status		X	X		
Mother's Education FE			X		
Family FE				X	
Family x Year FE					X

Table 5B: This table shows the estimates of a linear regression of test scores in mathematics (Panel A) and reading (Panel B) standardized by year and grade on the cumulative exposure to foreign-born students, computed as the average share of foreign-born students across previous school-specific cohorts including the current grade, and several controls. All regressions are run on an unbalanced longitudinal sample of Black U.S.-born students, observed in grades from 3rd to 10th, who speak English at home and have at least one sibling, using observations in academic years in which at least two students are observed for each family. Individual controls include: gender, age in months, special education, and birth order fixed effects. Lunch status is a dummy variable equal to 1 if the student is eligible for free or reduced-price lunch. Mother's education fixed effects are three dummy variables equal to 1 if the mother of the student has a high school diploma, some college, or a four-year college or more, respectively. Robust standard errors in parenthesis clustered by school-cohort. Beta standardized coefficients in squared parenthesis below standard errors. *** p<0.01, ** p<0.05, * p<0.1.

	Standardized scores (3rd-10th grade)				
	Sample restriction: Lunch Status = No Free/Reduced-price				
	(1)	(2)	(3)	(4)	(5)
	Panel A: Math standardized score				
Foreign-born Exposure	-0.460***	-0.424***	-0.296***	-0.002	-0.034
	(0.067)	(0.065)	(0.061)	(0.080)	(0.113)
	[-0.028]	[-0.025]	[-0.018]	[0.000]	[-0.002]
Observations	611,698	611,698	610,918	611,698	611,698
R-squared	0.218	0.235	0.270	0.672	0.763
Mean LHS	0.475	0.475	0.475	0.475	0.475
SD LHS	0.867	0.867	0.867	0.867	0.867
	Panel B: Reading standardized score				
Foreign-born Exposure	-0.478***	-0.449***	-0.309***	-0.091	-0.194*
	(0.064)	(0.063)	(0.059)	(0.079)	(0.106)
	[-0.028]	[-0.026]	[-0.018]	[-0.005]	[-0.011]
Observations	658,656	658,656	657,839	658,656	658,656
R-squared	0.197	0.210	0.244	0.636	0.733
Mean LHS	0.459	0.459	0.459	0.459	0.459
SD LHS	0.889	0.889	0.889	0.889	0.889
Individual Controls	X	X	X	X	X
School x Year FE	X	X	X	X	X
Grade x Year FE	X	X	X	X	X
Race FE		X	X		
Lunch Status		X	X		
Mother's Education FE			X		
Family FE				X	
Family x Year FE					X

Table 6A: This table shows the estimates of a linear regression of test scores in mathematics (Panel A) and reading (Panel B) standardized by year and grade on the cumulative exposure to foreign-born students, computed as the average share of foreign-born students across previous school-specific cohorts including the current grade, and several controls. All regressions are run on an unbalanced longitudinal sample of U.S.-born students not eligible for free or reduced-price lunch, observed in grades from 3rd to 10th, who speak English at home and have at least one sibling, using observations in academic years in which at least two students are observed for each family. Individual controls include: gender, age in months, special education, and birth order fixed effects. Lunch status is a dummy variable equal to 1 if the student is eligible for free or reduced-price lunch. Mother's education fixed effects are three dummy variables equal to 1 if the mother of the student has a high school diploma, some college, or a four-year college or more, respectively. Robust standard errors in parenthesis clustered by school-cohort. Beta standardized coefficients in squared parenthesis below standard errors. *** p<0.01, ** p<0.05, * p<0.1

	Standardized scores (3rd-10th grade)				
	Sample restriction: Lunch Status = Free/Reduced-price				
	(1)	(2)	(3)	(4)	(5)
	Panel A: Math standardized score (3rd-10th grade)				
Foreign-born Exposure	0.368***	0.283***	0.301***	0.452***	0.399***
	(0.053)	(0.050)	(0.049)	(0.074)	(0.102)
	[0.020]	[0.016]	[0.017]	[0.025]	[0.022]
Observations	735,588	735, 588	733,623	735, 588	735, 588
R-squared	0.250	0.280	0.293	0.620	0.728
Mean LHS	-0.303	-0.303	-0.302	-0.303	-0.303
SD LHS	0.952	0.952	0.952	0.952	0.952
	Panel B: Reading standardized score				
Foreign-born Exposure	0.312***	0.243***	0.267***	0.356***	0.300***
	(0.048)	(0.045)	(0.044)	(0.064)	(0.085)
	[0.018]	[0.014]	[0.015]	[0.020]	[0.017]
Observations	791,482	791, 482	789,439	791, 482	791, 482
R-squared	0.267	0.293	0.307	0.615	0.716
Mean LHS	-0.319	-0.319	-0.319	-0.319	-0.319
SD LHS	0.932	0.932	0.932	0.932	0.932
Individual Controls	X	X	X	X	X
School x Year FE	X	X	X	X	X
Grade x Year FE	X	X	X	X	X
Race FE		X	X		
Lunch Status		X	X		
Mother's Education FE			X		
Family FE				X	
Family x Year FE					X

Table 6B: This table shows the estimates of a linear regression of test scores in mathematics (Panel A) and reading (Panel B) standardized by year and grade on the cumulative exposure to foreign-born students, computed as the average share of foreign-born students across previous school-specific cohorts including the current grade, and several controls. All regressions are run on an unbalanced longitudinal sample of U.S.-born students eligible for free or reduced-price lunch, observed in grades from 3rd to 10th, who speak English at home and have at least one sibling, using observations in academic years in which at least two students are observed for each family. Individual controls include: gender, age in months, special education, and birth order fixed effects. Lunch status is a dummy variable equal to 1 if the student is eligible for free or reduced-price lunch. Mother's education fixed effects are three dummy variables equal to 1 if the mother of the student has a high school diploma, some college, or a four-year college or more, respectively. Robust standard errors in parenthesis clustered by school-cohort. Beta standardized coefficients in squared parenthesis below standard errors. *** p<0.01, ** p<0.05, * p<0.1.

	Standardized scores (3rd-10th grade)		
	Sample restriction: siblings who go to the same initial school		
	(1)	(2)	(3)
	IV	*Red. Form*	*OLS*
	Panel A: Math standardized score		
Foreign-born Exposure	0.319**		0.338***
	(0.155)		(0.068)
	[0.018]		[0.019]
Foreign-born Exposure (Predicted)		0.139**	
		(0.067)	
		[0.008]	
Observations	821,892	821,892	821,892
R-squared	-	0.668	0.668
Mean LHS	0.192	0.192	0.192
SD LHS	0.954	0.954	0.954
First stage (coefficient)	0.434***	-	-
First stage (se)	0.009	-	-
First stage (F stat)	2,274	-	-
	Panel B: Reading standardized score		
Foreign-born Exposure	0.520***		0.322***
	(0.144)		(0.063)
	[0.030]		[0.018]
Foreign-born Exposure (Predicted)		0.233***	
		(0.065)	
		[0.013]	
Observations	880,812	880,812	880,812
R-squared	-	0.654	0.654
Mean LHS	0.169	0.169	0.169
SD LHS	0.962	0.962	0.962
Individual controls	X	X	X
Year x Grade FE	X	X	X
Family x Initial School FE	X	X	X
First stage (coefficient)	0.448***	-	-
First stage (se)	0.009	-	-
First stage (F stat)	2,454	-	-

Table 7: This table shows results on the first instrumental variable approach described in the text. Column (1) presents the Two Stage Least Square coefficient, Column (2) presents the reduced form coefficient, and Column (3) shows the OLS version of the coefficient. The construction of the predicted Foreign-born exposure is described in the text. All regressions are run on an unbalanced longitudinal sample of U.S.-born students observed in grades from 3rd to 10th, who speak English at home and have at least one sibling. The sample is further restricted to students from families where all siblings attended the same initial school (i.e. the first school a student is observed in). Individual controls include: gender, age in months, special education, birth order fixed effects. Year x grade fixed effects are indicators for each unique year-grade combination. Family x Initial school fixed effects are indicators for each unique family-initial school combination. Robust standard errors in parenthesis clustered by school-cohort. Beta standardized coefficients in squared parenthesis below standard errors. *** p<0.01, ** p<0.05, * p<0.1.

	Sample: 3rd to 10th grade			
	(1)	(2)	(3)	(4)
<u>Outcomes:</u>	*Math stdz score*	*Reading stdz score*	*Math stdz score*	*Reading stdz score*
Foreign-born Exposure	0.335***	0.176**		
	(0.080)	(0.069)		
	[0.018]	[0.009]		
Foreign-born Exposure (low SES)			0.355***	0.193**
			(0.105)	(0.090)
			[0.014]	[0.008]
Foreign-born Exposure (high SES)			0.300**	0.147
			(0.139)	(0.126)
			[0.008]	[0.004]
Black Exposure	-0.013	-0.119***	-0.012	-0.118***
	(0.044)	(0.038)	(0.044)	(0.038)
	[-0.003]	[-0.028]	[-0.003]	[-0.028]
White Exposure	0.058	0.010	0.059	0.011
	(0.046)	(0.041)	(0.046)	(0.041)
	[0.015]	[0.003]	[0.015]	[0.003]
Asian Exposure	0.377***	0.467***	0.380***	0.470***
	(0.132)	(0.121)	(0.133)	(0.121)
	[0.008]	[0.010]	[0.008]	[0.010]
Free or Reduced-price Lunch Exposure	-0.232***	-0.213***	-0.235***	-0.216***
	(0.033)	(0.030)	(0.035)	(0.031)
	[-0.055]	[-0.050]	[-0.055]	[-0.050]
Limited English Proficiency Exposure	-0.109**	-0.167***	-0.110**	-0.168***
	(0.052)	(0.046)	(0.052)	(0.046)
	[-0.009]	[-0.013]	[-0.009]	[-0.013]
Special Education Exposure	-0.195***	-0.262***	-0.195***	-0.263***
	(0.047)	(0.040)	(0.047)	(0.040)
	[-0.010]	[-0.013]	[-0.010]	[-0.013]
Individual Controls	X	X	X	X
School x Year FE	X	X	X	X
Grade x Year FE	X	X	X	X
Family x Year FE	X	X	X	X
Observations	1,347,286	1,450,138	1,347,286	1,450,138
R-squared	0.769	0.752	0.769	0.752
Mean LHS	0.0504	0.0340	0.0504	0.0340
SD LHS	0.993	0.992	0.993	0.992

Table 8: This table shows estimates from models equivalent to those reported in Column (5) of Table 4, adding controls for other exposures. These cumulative exposure variables are calculated following equation 1 in the text and computed as leave-out-means. Robust standard errors in parenthesis clustered by school-cohort. Beta standardized coefficients in squared parenthesis below standard errors. *** p<0.01, ** p<0.05, * p<0.1.

	Standardized scores (3rd-10th grade)				
	(1)	(2)	(3)	(4)	(5)
Restriction:	*Full sample*	*No free lunch*	*Free lunch*	*White*	*Black*
	Panel A: Math standardized score				
Foreign-born Exposure	0.222***	-0.038	0.382***	0.143	0.403***
	(0.078)	(0.118)	(0.108)	(0.112)	(0.144)
	[0.012]	[-0.002]	[0.021]	[0.007]	[0.022]
Immigrant performance index	0.037***	0.031**	0.037***	0.032***	0.039**
(Math score)	(0.008)	(0.013)	(0.011)	(0.011)	(0.017)
	[0.010]	[0.009]	[0.011]	[0.009]	[0.011]
Observations	1,271,246	585,107	686,139	764,962	374,307
R-squared	0.777	0.770	0.740	0.774	0.730
Mean LHS	0.0579	0.481	-0.301	0.314	-0.489
SD LHS	0.993	0.867	0.952	0.911	0.951
	Panel B: Reading standardized score				
Foreign-born Exposure	0.138**	-0.171	0.323***	0.048	0.326***
	(0.067)	(0.110)	(0.089)	(0.104)	(0.115)
	[0.007]	[-0.010]	[0.018]	[0.002]	[0.019]
Immigrant performance index	0.044***	0.050***	0.033***	0.036***	0.034**
(Reading score)	(0.008)	(0.012)	(0.010)	(0.010)	(0.015)
	[0.012]	[0.013]	[0.009]	[0.010]	[0.009]
Observations	1,371,517	630,822	740,695	824,567	405,141
R-squared	0.760	0.740	0.728	0.749	0.719
Mean LHS	0.0401	0.463	-0.318	0.296	-0.507
SD LHS	0.992	0.889	0.932	0.933	0.904
Individual Controls	X	X	X	X	X
School x Year FE	X	X	X	X	X
Grade x Year FE	X	X	X	X	X
Family x Year FE	X	X	X	X	X

Table 9: This table shows the estimates of a linear regression of test scores in mathematics (Panel A) and reading (Panel B) standardized by year and grade on the cumulative exposure to foreign-born students, computed as the average share of foreign-born students across previous school-specific cohorts including the current grade. The regression also includes a school-cohort index of foreign-born performance computed as a weighted average of country-specific mean math and reading test scores, weighted by the share of students from a given country, in a given school-specific cohort. All columns show estimates from models equivalent to the one reported in Column (5) of Table 4 performed on different sub-samples. Robust standard errors in parenthesis clustered by school-cohort. Beta standardized coefficients in squared parenthesis below standard errors. *** p<0.01, ** p<0.05, * p<0.1.

	Standardized scores (3rd-10th grade)				
	(1)	(2)	(3)	(4)	(5)
Restriction:	*Full sample*	*No free lunch*	*Free lunch*	*White*	*Black*
			Panel A: Math standardized score		
Foreign-born Exposure	0.209***	-0.040	0.376***	0.142	0.397***
	(0.077)	(0.118)	(0.108)	(0.112)	(0.144)
	[0.011]	[-0.002]	[0.021]	[0.007]	[0.021]
Immigrant performance index (Behavior)	-0.222***	-0.217**	-0.285***	-0.206**	-0.267**
	(0.059)	(0.107)	(0.092)	(0.092)	(0.128)
	[-0.008]	[-0.007]	[-0.010]	[-0.006]	[-0.010]
Observations	1,271,246	585,107	686,139	764,962	374,307
R-squared	0.777	0.770	0.740	0.774	0.730
Mean LHS	0.0579	0.481	-0.301	0.314	-0.489
SD LHS	0.993	0.867	0.952	0.911	0.951
			Panel B: Reading standardized score		
Foreign-born Exposure	0.112*	-0.177	0.315***	0.043	0.317***
	(0.067)	(0.110)	(0.089)	(0.104)	(0.115)
	[0.006]	[-0.010]	[0.018]	[0.002]	[0.018]
Immigrant performance index (Behavior)	-0.255***	-0.350***	-0.332***	-0.264***	-0.315***
	(0.054)	(0.100)	(0.080)	(0.084)	(0.107)
	[-0.009]	[-0.010]	[-0.012]	[-0.008]	[-0.012]
Observations	1,371,517	630,822	740,695	824,567	405,141
R-squared	0.760	0.740	0.727	0.749	0.720
Mean LHS	0.0401	0.463	-0.319	0.296	-0.507
SD LHS	0.992	0.889	0.932	0.933	0.904
Individual Controls	X	X	X	X	X
School x Year FE	X	X	X	X	X
Grade x Year FE	X	X	X	X	X
Family x Year FE	X	X	X	X	X

Table 10: This table shows the estimates of a linear regression of test scores in mathematics (Panel A) and reading (Panel B) standardized by year and grade on the cumulative exposure to foreign-born students, computed as the average share of foreign-born students across previous school-specific cohorts including the current grade. The regression also includes a school-cohort index of foreign-born behavioral performance, computed as a weighted average of country-specific average likelihood of being involved in a disciplinary incident, weighted by the share of students from a given country, in a given school-specific cohort. All columns show estimates from models equivalent to the one reported in Column (5) of Table 4 performed on different sub-samples. Robust standard errors in parenthesis clustered by school-cohort. Beta standardized coefficients in squared parenthesis below standard errors. *** p<0.01, ** p<0.05, * p<0.1.

	Math Score (Standardized)		Reading Score (Standardized)		Incidents (indicator)	
	Obs.	Mean	Obs.	Mean	Obs.	Mean
U.S.-born speaking English (Whole sample)	1,347,287	0.05	1,450,139	0.034	1,450,139	0.169
Immigrants who go to school with those above	58,736	0.006	60,663	-0.071	60,663	0.119
White U.S.-born speaking English	811,790	0.305	873,281	0.288	873,281	0.105
Immigrants who go to school with those above	49,496	0.093	51,101	0.026	51,101	0.11
Black U.S.-born speaking English	399,586	-0.495	430,975	-0.511	430,975	0.31
Immigrants who go to school with those above	45,497	-0.180	47,243	-0.275	47,243	0.142
No-FRPL U.S.-born speaking English	611,698	0.475	658,656	0.459	658,656	0.074
Immigrants who go to school with those above	49,168	0.170	50,805	0.101	50,805	0.104
FRPL U.S.-born speaking English	735,589	-0.303	791,483	-0.319	791,483	0.247
Immigrants who go to school with those above	54,774	-0.137	56,653	-0.22	56,653	0.131

Table 11: This table shows descriptive statistics of test scores and incident rates across different subset of students. It shows the mean of each variable for the sample of U.S.-born students speaking English, and for the foreign-born students who are in the same school-cohort. These statistics are shown first for the entire sample of U.S.-born students and then for four different subsamples, based on reported race/ethnicity and free lunch eligibility.

	Standardized scores (3rd-10th grade)				
	(1)	(2)	(3)	(4)	(5)
Restriction:	*Full sample*	*No free lunch*	*Free lunch*	*White*	*Black*
	Panel A: Math standardized score				
Foreign-born Exposure	0.641***	0.285	0.953***	0.506**	1.038***
(LTO above US)	(0.167)	(0.220)	(0.264)	(0.208)	(0.370)
	[0.010]	[0.006]	[0.014]	[0.009]	[0.015]
Foreign-born Exposure	0.203*	-0.012	0.298*	0.109	0.430*
(LTO below US)	(0.123)	(0.184)	(0.174)	(0.178)	(0.246)
	[0.006]	[-0.000]	[0.009]	[0.003]	[0.012]
Immigrant performance index	0.028***	0.022*	0.028**	0.024**	0.027
(Math score)	(0.009)	(0.013)	(0.012)	(0.011)	(0.018)
	[0.008]	[0.006]	[0.008]	[0.007]	[0.008]
Observations	1,271,246	585,107	686,139	764,962	374,307
R-squared	0.778	0.770	0.740	0.774	0.730
Mean LHS	0.0592	0.481	-0.301	0.314	-0.489
SD LHS	0.993	0.867	0.952	0.911	0.951
	Panel B: Reading standardized score				
Foreign-born Exposure	0.457***	0.105	0.669***	0.254	0.841***
(LTO above US)	(0.152)	(0.208)	(0.230)	(0.196)	(0.311)
	[0.007]	[0.002]	[0.010]	[0.005]	[0.013]
Foreign-born Exposure	0.178	-0.288*	0.424***	-0.036	0.552***
(LTO below US)	(0.108)	(0.173)	(0.147)	(0.164)	(0.204)
	[0.005]	[-0.010]	[0.013]	[-0.001]	[0.016]
Immigrant performance index	0.038***	0.043***	0.027**	0.032***	0.024
(Reading score)	(0.008)	(0.013)	(0.011)	(0.011)	(0.015)
	[0.010]	[0.012]	[0.008]	[0.009]	[0.007]
Observations	1,371,517	630,822	740,695	824,567	405,141
R-squared	0.761	0.740	0.728	0.749	0.720
Mean LHS	0.0414	0.463	-0.318	0.296	-0.507
SD LHS	0.992	0.889	0.932	0.933	0.904
Individual Controls	X	X	X	X	X
School x Year FE	X	X	X	X	X
Grade x Year FE	X	X	X	X	X
Family x Year FE	X	X	X	X	X

Table 12: This table shows the estimates of a linear regression of test scores in mathematics (Panel A) and reading (Panel B) standardized by year and grade on the cumulative exposure to foreign-born students recalculated splitting migrants by the LTO (associated with their country of origin), above and below US LTO. Since LTO data for Cuba, Jamaica, and Haiti are unavailable, we also control for a residual exposure component for these three countries (not shown in the Table). All columns show estimates from models equivalent to the one reported in Column (5) of Table 4 performed on different sub-samples. Robust standard errors in parenthesis clustered by school-cohort. Beta standardized coefficients in squared parenthesis below standard errors. *** p<0.01, ** p<0.05, * p<0.1.