DISCUSSION PAPER SERIES

DP15933

Diversity in Schools: Immigrants and the Educational Performance of U.S. Born Students

David Figlio, Paola Giuliano, Riccardo Marchingiglio, Umut Ozek and Paola Sapienza

LABOUR ECONOMICS

CEPR

ISSN 0265-8003

Diversity in Schools: Immigrants and the Educational Performance of U.S. Born Students

David Figlio, Paola Giuliano, Riccardo Marchingiglio, Umut Ozek and Paola Sapienza

Discussion Paper DP15933
Published 16 March 2021
Submitted 16 March 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Labour Economics

Diversity in Schools: Immigrants and the Educational Performance of U.S. Born Students

Abstract

We study the effect of exposure to immigrants on the educational outcomes of US-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 US-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 to partial out the unobserved non-random selection into schools, and find that the presence of immigrant students has a positive effect on the academic achievement of US-born students, especially for students from disadvantaged backgrounds. Moreover, the presence of immigrants does not affect negatively the performance of affluent US-born students, who typically show a higher academic achievement compared to immigrant students. We provide suggestive evidence on potential channels.

JEL Classification: I21, I24, J15

Keywords: Immigrant students, Educational Attainment

David Figlio - figlio@northwestern.edu
Northwestern University

Paola Giuliano - paola.giuliano@anderson.ucla.edu
UCLA and CEPR

Riccardo Marchingiglio - riccardo.marchingiglio@analysisgroup.com
Analysis Group

Umut Ozek - uozek@air.org
AIR

Paola Sapienza - paola-sapienza@kellogg.northwestern.edu
Northwestern University

Acknowledgements
This research was funded by the Russell Sage Foundation. We thank Jennifer Hunt, Andrea Ichino, Victor Lavy, Daniele Paserman, Bruce Sacerdote, Cynthia van der Werf, and participants at various seminars and conferences for helpful comments, and Gaia Dossi for truly outstanding research assistantship. We are especially grateful to the Florida Department 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.

Diversity in Schools:

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

David Figlio, Northwestern University and NBER

Paola Giuliano, UCLA and NBER

Riccardo Marchingiglio, Northwestern University

Umut Ozek, American Institute for Research

Paola Sapienza, Northwestern University and NBER

Current Draft: March 2021

Abstract

We study the effect of exposure to immigrants on the educational outcomes of US-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 US-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 to partial out the unobserved non-random selection into schools, and find that the presence of immigrant students has a positive effect on the academic achievement of US-born students, especially for students from disadvantaged backgrounds. Moreover, the presence of immigrants does not affect negatively the performance of affluent US-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] This research was funded by the Russell Sage Foundation. We thank Jennifer Hunt, Andrea Ichino, Victor Lavy, Daniele Paserman, Bruce Sacerdote, Cynthia van der Werf, and participants at various seminars and conferences for helpful comments, and Gaia Dossi for truly outstanding research assistantship. We are especially grateful to the Florida Department 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.

1. Introduction

Over the past 50 years, immigration rates into the United States have risen dramatically. As a result, almost one out of four (23 percent) public school students in the United States came from an immigrant household in 2015 (either foreign-born students or second generation students), with concentrations over 70% in several school districts (and even higher in some communities, including as high as 93 percent in northeast Miami-Dade County, Florida, 91 percent in Jackson Heights and North Corona, New York, and 85 percent in Westpark Tollway, Texas). 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 US-born student population.

Given the sheer size of immigrants in US schools and their unique cultural backgrounds, studying their impact on US-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 US-born students (Fix and Zimmerman, 1993). On the other hand, especially some groups of immigrants, through hard work and resilience, outperform non-immigrant students with similar socio-economic backgrounds (Hsin and Xie, 2014) and can positively affect exposed US-born students' attitudes and behavior (Hunt, 2016). Yet, the academic research about the impact of immigrant students on the educational performance of US-born students is limited[2], especially compared to the literature about the effect of immigrants on labor market outcomes[3], primarily due to two important empirical challenges.

First, immigrant students are not randomly assigned to schools, and are more likely to enroll in schools educating students from disadvantaged backgrounds (e.g., Card 2001; Figlio and Ozek

[2] There are a few studies that examine the effects of immigrant students on US-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 US-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 US-born students. There are also studies in the US context that investigate the effects of specific immigrant groups (e.g., refugees) on US-born student outcomes (Figlio and Ozek, 2019; Morales 2020; Ozek, forthcoming; Van der Werf 2021).

[3] The results on the labor market effects of immigrants are mixed. Friedberg and Hunt (1995) and Card (2001) find that immigration has a small negative effect on employment rates and wages of low-skilled native-born workers, whereas Borjas (2003) finds a larger negative impact on the wages of low-skilled native-born workers, particularly those without high school diplomas. Ottaviano and Peri (2012) argue that the small negative impact of immigration on low-skilled native-born wages and employment is outweighed by increased demand for higher-level workers and improved physical capital.

2019). Second, US-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 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 of these factors imply that immigrant exposure is negatively correlated with the SES of US-born students. Therefore, research that does not address the non-random selection of US-born students is likely to estimate a correlation between immigrant exposure and US-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 US-born students using unique administrative data from Florida that link population-level school records and birth records. There are several advantages in using this dataset. First, 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. Second, 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 to address the possibility that families select schools differentially for each child using an instrumental variable approach. Specifically, we build a measure of predicted immigrant exposure using aggregate school-to-school transition probabilities, for each kid at each subsequent grade, starting from the first grade the student is observed. For example, 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.

[4] Several papers have studied this topic outside the US finding 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.

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 run this regression, 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 US-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 US-born student selection into schools is that there is strong evidence that US-born students indeed sort into schools on the basis of immigrant concentration. 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% and 13%, respectively) increases the score in mathematics and reading by 2.7% and 1.7% of a standard deviation, respectively. This effect corresponds to 8.5% 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-priced lunch (FRPL) eligible students and for Black students.

These heterogeneous effects hint at the possibility that US-born students with different demographics are exposed to immigrant students with different levels of academic performance. Indeed, we find that immigrants who go to schools with higher shares of Black or FRPL eligible students underperform immigrants who go to school with higher shares of White or FRPL ineligible US-born students. On the other hand, immigrants going to high-Black or high-poverty schools outperform US-born students in the same school-year-grade, while White and FRPL ineligible US-born students, on average, outperform foreign-born students in the same school-year-grade. We also find that immigrants have fewer disciplinary incidents than US-born students, but the difference in behavior is mostly observed in schools where the majority of US-born are disadvantaged students.

If academic achievement and behavior matter to explain the positive effect of immigrant exposure, one should expect that the higher the fraction of high performing and "better behaved" immigrants, the more positive the impact on US-born students. The reflection problem (Manski, 1993) and endogeneity issues do not allow the identification of the causal impact of the achievement of immigrants on the performance of US-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 and/or disciplinary behavior 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 proxies for the *potential* academic achievement (and behavior) of immigrants. We show that the presence of immigrants with higher expected academic performance correlates with better scores of US-born students in the overall sample and across every subsample. After controlling for expected academic achievement of the immigrants, the coefficient on the immigrant exposure variable remains similar to our baseline specification.

It is also possible that what matters is not just the absolute performance of the immigrants, but how much better or worse they perform vis-à-vis US-born students in the same school-grade-year. We cannot test this hypothesis directly because we do not observe the potential performance of each US-born student in absence of their exposure to immigrants. Nonetheless, we find suggestive evidence that the effect on US-born students is larger when the immigrants systematically outperform US-born students. Overall, these results suggest that immigrant students do not affect negatively US-born students, even when the immigrants' academic achievement is lower than the US-born students, and may have a positive impact on US-born students when immigrants outperform them.

2. Data and Variables of Interest

2.1 Data Sources

We use a unique 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 contain information on K-12 students who attended Florida public schools between 2002-2003 and 2011-2012. The data contain 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.[5] 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 US-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). Therefore, given our definition, *immigrant* is equivalent to *foreign-born*.[6] Because we do not have birth certificates for Puerto Rican students, we are unable to include them in the sample of US-born students. Therefore, the best way to treat Puerto Ricans in this analysis is not obvious: we can either include them with the foreign students to calculate the foreign exposure measure or we can exclude them altogether from the analysis. We adopt both strategies. In the baseline regression, we treat Puerto Ricans as "immigrants" on the ground that they are culturally distinct from many 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 immigrants children born in the US with parents born abroad

[5] 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% 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.

[6] 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 can at least 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; results remain highly consistent regardless of whether we include or exclude these military-intensive counties (at request by the authors).

(second generation immigrants). Because we do not have information on the immigrant status of the father we do not follow this strategy.

2.3 Measure of Immigrant Exposure

We adopt a cumulative measure of immigrant exposure, in which we aggregate the share of foreign-born students to whom a US-born student has been exposed from kindergarten to the time of observation (measured at the school-grade-year cell level.) This is the most flexible approach to studying exposure because it does not require us to take a stand regarding the degree to which the effects of immigrant exposure persist beyond contemporaneous exposure. 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, we consider a general model of immigrant exposure 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:

$$E_{isgt} = \sum_{g' \leq g} IMMIGRANT\_SHARE_{isg't} * e^{\left(1-\left(\lambda * (g-g')\right)\right)} \Big/ \sum_{g' \leq g} e^{\left(1-\left(\lambda * (g-g')\right)\right)} \quad (1)$$

where $IMMIGRANT\_SHARE_{g'}$ is the exposure in grade g'.

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 peer students. This specification permits a wide range of models, from a model in which last year's exposure is just as influential as contemporaneous exposure ($\lambda = 0$) to a model in which only contemporaneous exposure matters (λ increasing to infinity). In our baseline, we begin with a zero decay model (λ=0) and later we expand it to include different λs. Thus, our baseline definition of cumulative exposure for each student *i* in school *s*, grade *g* and academic year *t,* is:

$$Immigrant\ Exposure_{isgt} = \frac{1}{g} \sum_{g' \leq g} IMMIGRANT\_SHARE_{isg't} \quad (2)$$

2.4 Outcome Variables

Our main outcomes of interest are Florida Comprehensive Assessment Test (FCAT) scores in mathematics and reading from grade 3 to grade 10 (the first and last year of statewide testing).

Because Florida transitioned to a new 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 dummies), a measure of low-income status (a dummy for whether the student is eligible to receive free-or-reduced-priced 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).[7]

2.6 Definition of US-born Students and Construction of the Sample of Interest

We define as *US-born students* all students born in the US 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.[8] Given our identification strategy, in our analysis we select the sample for which (1) we have test scores and (2)

---

[7] The race/ethnicity variables are collected by 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"). FLDOE forces a choice between White, Black, or Hispanic, so each student chooses a single identity. To qualify for free or reduced lunch, the family income has to be respectively below 185% and 130% 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.

[8] In Table A2 of the On-line Appendix, we report the descriptive statistics of the US-born students going to private and public schools in Florida. Using Census 2000 data, we compare the population of immigrant students attending public schools in Florida (93%) with those of the US-born (88%). US-born students, on average, are exposed to immigrant children who have lower SES than themselves, independently from the school setting: the family income of US-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.

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) observations for reading (math) scores.[9] The US-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 US-born students speaking English at home (Table A1.A), we obtain 6,341,333 observations (columns 7 to 9). From this sample, restricting to observations in families with at least two children in school in a given academic year leads to 1,450,139 observations for reading scores and 1,347,287 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.

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 U.S. 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 US-born students (0.044 and 0.052). Immigrants are also poorer (68% are FRPL eligible) than US-born (54%) and vary significantly in terms of racial background, language ability, and academic performance. In terms of racial composition, most of them are Hispanic (61%), while among US-born students only 22% are Hispanic. Immigrants are also more exposed to other immigrants (18% compared to US-born students who are exposed to 8% of immigrants). Consistent with evidence in other domains where immigrants tend to commit fewer crimes than non-immigrants (Nunn et al., 2018), immigrant students are involved in fewer disciplinary incidents (0.121) than US-born students (0.137).

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

[9] The discrepancy between reading and math observations is due to the fact that 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 US-born Students' Exposure to Immigrant Students

In the sample used in our regressions, students have an average cumulative exposure to immigrant students of 6%, but there is a lot of variation across schools. Figure 1 shows the distribution of the fraction of immigrants by institution, grade, and year. Most schools tend to have a fraction of immigrants lower than 10%; however, there is a non-trivial number of schools with a fraction of immigrants larger than 20%. 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 concentration of exposure for US-born students, across grades (Figure 4A) and by academic year (Figure 4B). The average concentration by academic year appears to be stable, suggesting that there is not an increase over time in cohorts of immigrants. Instead, Figure 4A shows that from grade 3 to 10, there is an increase in the fraction of immigrants, either because many first generation immigrants enroll in schools in higher grades (after immigrating) and/or because of lower dropout rates of immigrants in higher grades.

We then look at whether US-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 US-born students by race (Figure 5) and we observe a substantial gap in exposure to foreign-born. White students experience the lowest exposure to immigrants (around 6%), Hispanic students the largest (around 12%), and Black students somewhere in between (8%). Also, FRPL eligible students see a larger fraction of immigrants than non-eligible students but the difference is less pronounced than when we split the sample by race (Figure 6).

Table 2 lists the top 10 countries of origin of immigrants in Florida facing our sample of US-born students and facing the sub-samples of US-born divided by race.[10] The top 10 countries of origin in the overall sample are all Latin American countries. Together they constitute 65% of the immigrant

[10] Note that, as mentioned above, we consider models in which we treat Puerto Ricans, all of whom are US citizens, either as "US-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.

sample. For the school-specific cohorts where the majority of US-born students is White, Hispanic, or Black the results vary.[11] In school-cohorts where the majority of US-born students are White, Mexico represents the largest fraction (13%); in addition, several non-Latin American countries are at the top of the distribution: Germany (5%), Canada (4%) and China (3%). In school-cohorts where the majority of US-born students are Hispanic, immigrants come mostly from Latin American countries, especially Cuba (46%). The 10 largest countries of origin represent 85% of the overall immigrant distribution. Finally, in school-specific cohorts where the majority of US-born students are Black, the largest fraction of immigrants comes from Haiti (41%) and Jamaica (13%), and 78% of the immigrant exposure comes from 10 countries.

In Table 3, we divide again the set of school-specific cohorts by the predominant ethnicity of the US-born student sample and examine the racial/ethnic composition of the immigrants. In the overall sample, US-born students are exposed to Hispanic immigrants (62%), followed by Black immigrants (17%), and White immigrants (13%). However, there is a large heterogeneity in exposure once we split the schools by predominant races/ethnicities. US-born students in predominantly White schools are exposed to fewer Hispanic immigrants (46%), while White students are 29% and Asian students 13%. Students attending predominantly Hispanic schools are exposed to 92% of immigrants of Hispanic origin and a smaller fraction of Black and White students (each group constitutes only 3% of the total immigrant distributions). Finally, students going to predominantly Black schools are exposed to mostly Black immigrants (63%), followed by Hispanic (28%) and Asian immigrants (only 5%). Taken all together, these initial descriptive statistics show that US-born students are exposed to different subgroups of immigrant peers, depending on their race and ethnicity.

3. Empirical Analysis

3.1 Main Results

Tables 4 and 5 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

[11] 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 61,836 school-specific cohorts with at least one foreign-born student, 27,067 school-specific cohorts have a majority of White US-born students, 8,336 school-specific cohorts have a majority of Black US-born students, while 6,326 school-specific cohorts have a majority of Hispanic US-born students. The remaining schools in our sample have either a foreign-born majority or a US-born majority of another racial/ethnic group.

year t on $Immigrant\ Exposure_{isgt}$ defined in equation (2).[12] Our most demanding specification is the following:

$$Y_{isgt} = \alpha + Immigrant\ Exposure_{isgt}\beta + Z'_{isgt}\gamma + \delta_{gt} + \vartheta_{st} + \phi_{ft} + \varepsilon_{isgt} \qquad (3)$$

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. We cluster the standard errors at the cohort-school level.

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 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 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 US-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.[13] Our specification provides an effective way to control for selection into schools by families to the extent that this

[12] 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 US-born students, mechanically increasing their score), in robustness analysis we repeat the same specification and standardize scores using only our sample of US-born students. The results are substantially the same.

[13] As we include family fixed effects, we remove the controls for race, lunch status, and mother's education.

selection is made at the family level and it is not done differentially for each child.[14] 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.

To illustrate how much variation is captured by the different fixed effects,[15] 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 fixed-effects capture a good part of the variation, a lot still remains to be explained, indicating that we are running a meaningful model. 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 parameters.

To understand the economic magnitude of our effects, we compare our estimates to the relationship between maternal education and student outcomes. The beta coefficient (0.0121) of immigrant exposure in Column 5 is equal, for mathematics, to 8.5% of the difference in standardized test scores between students whose mothers does not have a high school diploma and students whose mother has an high school diploma (the beta coefficient in this case is 0.143).[16] The beta coefficient of immigrant exposure on reading scores (0.0058) is lower than math and corresponds to 4% of the difference in standardized test scores between students whose mother does not have a high school

[14] We will return to the assumption that the family does not choose different schools for each child, based on their attitudes, later.

[15] In a recent working paper, Miller, Shenhav and Grosz (2019) 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 suffering from selection into identifcation. First, in Section 4.2 we will provide evidence that our treatment effects do not change based on differences in school choice between siblings. Second, after partialling out family fixed effects, the distributions of our treatment and outcome variables still show significant variation (Figures A1-A3).

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

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 10^{th} to the 90^{th} percentile in the distribution of cumulative exposure (1% and 13%, respectively) would increase the score in mathematics and reading by 2.8% and 1.7% of a standard deviation, respectively.[17] We also study whether these magnitudes are different across grades by plotting the coefficient of the immigrant share interacted with grade in the baseline specification (Figure A4). We do not find any significant statistical difference across grades.

As mentioned in Section 3.2, we elected to present as our baseline specification a measure of cumulative exposure in which last year's exposure is just as important as this year's exposure (a decay parameter $\lambda=0$). However, we have estimated models with a wide variety of decay parameters, 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; these results are presented in Figure A5.

Our results are also robust to four additional specifications. The first pertains the treatment of Puerto Rican students. As discussed previously, we re-run our specification, excluding Puerto Ricans from the immigrant groups. The results are reported in the Appendix Table A3 and A4 and are consistent with our previous sample's findings and the beta coefficients have similar magnitudes. The second robustness includes third grade scores as a control and recalculate the measure of immigrant exposure from third grade to the current grade. In this specification, the beta coefficient of immigrant exposure is virtually identical to the preferred specifications of columns 4 and 5 (Table A5). In the third robustness, we re-run the specifications of Table 4, using a different definition of US-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 A6 and A7). 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 US-born students is reflecting these socioeconomic characteristics of the immigrants, in Table A8, we present our fourth robustness where we saturate our model introducing a vector of cumulative exposures to racial minority groups, FRPL eligible peers,

[17] Lavy and Schlosser (2011) study the peer effects of female students on students' academic achievements in Israel. They find that a 20-percentage-points increase in the proportion of female students translates into 4-5% 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 peers roughly corresponds to a 5 percent increase in female share in Lavy and Schlosser (2011).

and to peers with limited English proficiency. Specifically, analogously to our main measure of exposure described in equation (2), we calculate cumulative exposures to groups of students based on additional characteristics: race (Black, Asian, and Hispanic) and fraction of FRPL eligible students, and with limited English proficiency. Even in this saturated specification, the coefficient on immigrant cumulative exposure remains statistically significant, with a similar magnitude.

Taken together, these results suggest the presence of a strong selection of US-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 US-born students, in Tables 6 and 7, we split the sample by race/ethnicity and SES. In Tables 6A and 6B, we divide the sample into White and Black students and examine their performance in mathematics.[18] The conditional correlation between immigrant exposure and the performance of White US-born students is very similar to Table 4: without the inclusion of any family control, it is negative and significant, but becomes positive and significant with the inclusion of family fixed effects (the inclusion of family-year fixed effects reduces significance and size of the beta coefficient). The results for Black students are very different: the conditional correlation between immigrants' exposure and performance is stable and positive, independently of the controls included in the analysis. These results are consistent with the existing literature on "white flight" that suggests that White families are more likely to select into schools with a low fraction of minority and immigrant students (e.g., Betts and Fairlie, 2003). On the contrary, US-born Black students do not select specifically into schools based on immigrant shares.

In Tables 7A and 7B, to further validate this interpretation, we separate higher and lower-SES US-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. Conversely, the results for lower-SES students show that this group does not suffer from self-selection issues, similarly to Black students, and the coefficient is positive and significant in every specification.[19]

[18] 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. The results are similar for reading (see Table A9 in the Appendix).

[19] Similar results are obtained for reading tests (Tables A10 in the Appendix).

Another important difference that emerges from Tables 6 and 7 is that the impact of immigrant exposure has differential effects on different subgroups. Compared with the overall sample, the effect of immigrant exposure is twice as large for Black and FRPL eligible students; while for White and FRPL ineligible students the effect is null and not significant.

3.2 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 20.1 versus 34 months. The results are presented in Table A11. Column 1 repeats our preferred specification with the sub-sample of all families with only two children, column 2 presents results with children going to the same school, and column 3 shows results with children going to different schools. If anything, the results seem to be stronger for the subsample of children going to the same school.

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 time frame 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% 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%, the fraction of White/Black students is 68%/22%. Maternal education is also higher for the students in this group: fewer students have mother who dropped out of high school (15%), while more students have mothers who completed 4 years of college (24%).[20]

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

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

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 whole 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:

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

[20] In this IV estimation, we ignore the families who sends their kids to different initial schools. 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 5th grade) at the time in which the younger sibling enrolls in first grade. Among these families, 69% chose the same identical first school for both siblings in grade 1; 24% 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 sibling, 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 attitudes. The remaining 7% go to a first school which is different from the contemporaneous school of the older sibling: this sample has much worse educational attainment and lower SES. Since selection into schools seems to be a prerogative of high-SES families, the exclusion of this group from this IV analysis does not seem problematic.

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 bin scatters for the first stage, based on the unconditional model and the model including family and initial school fixed effects are presented respectively in Figures 7A and 7B 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 8 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 coefficient of immigrant exposure in the OLS is positive and significant. The instrumental variable coefficient is almost identical confirming that school choice is mostly done at the family level.

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.[21] Our findings are consistent (Table A12 in the Appendix).

4. Heterogeneity

Tables 6 and 7 have shown that immigrant exposure effects are stronger for less affluent and Black US-born students. One possible reason is that US-born students with different demographic characteristics face immigrants that, in relative terms, perform better or worse than them. The relative standing of immigrant students may drive some of the heterogeneous results. We examine this

[21] 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.

hypothesis by analyzing our two academic achievement outcomes (math and reading scores) and study whether relative academic differences could be driving our results. In Table 9, we calculate the mean of these two outcome variables for US-born students in our sample and for the corresponding groups of immigrant students going to school with them. We repeat this exercise for different SES sub-samples. In the first two columns of Table 9, we calculate the average math standardized scores. For the overall sample, US-born students have an average math score of 0.05, while the performance of the immigrant students going to the same school is lower, at 0.006. However, when we consider high- and low-SES students, the relative performance of immigrant students is very different. Immigrants going to school with FRPL ineligible US-born students have an average math score of 0.17, substantially lower than the score of their US-born schoolmates (0.475), but much higher than the average math score of the immigrants going to school with the US-born, FRPL eligible students (-0.137). Yet, these latter immigrants outperform US-born, FRPL eligible students (-0.303). These patterns are also reflected in the relative differences between White and Black US-born students and the population of immigrants going to school with them. Reading scores show very similar trends. These results suggests that part of the effect of immigrant exposure may be driven by average relative differences in performance, especially for those groups of students whose immigrant schoolmates outperform them (Black and FRPL eligible students where the majority of the effect is concentrated).

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). As for academic performance, the disadvantaged US-born students, who benefit more from the presence of immigrants are, on average, exposed to students behaving better than them, suggesting that one of the potential mechanisms could be exposure to less disruptive students. On average, immigrants behave better (0.119) than their classmates (0.169). However, when we split the data into sub-samples based on SES, we find again strong evidence of selection. FRPL ineligible and White US–born students go to school with immigrants who behave better (respectively, 0.104 and 0.11) than the immigrants going to school with FRPL eligible and Black US-born students (respectively, 0.131 and 0.142). Also, White and FRPL ineligible US-born students have fewer disciplinary incidents (respectively, 0.105 and 0.074), on average, than the immigrants going to school with them, while Black and FRPL eligible US-born students have substantially more disciplinary incidents (respectively, 0.31 and 0.247) than their immigrant schoolmates.

To study whether the academic performance of immigrants drives our results, we need to address two major challenges. First, it is impossible to analyze the importance of the *relative* difference between US-born students and their immigrant schoolmates because we do not observe the potential performance of each US-born student in absence of their exposure to immigrant students. Second, 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 US-born students. To address the latter problem, we can 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 (Figlio et al., 2019). We use these measures of expected academic performance to weight our cumulative exposure to immigrants and add this immigrant performance index to our baseline regressions.[22]

Table 10 shows the results for math scores when we include this immigrant performance index into our analysis. In our preferred specification, with the inclusion of family-year fixed effects (column 5), this weighted index has a positive and significant coefficient, with a very similar economic magnitude to the immigrant exposure's coefficient (0.0113). The size of the immigrant exposure coefficient does not change compared to our baseline specification.[23]

[22] 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 US-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 A6, Panel A in the Appendix) confirm large differences among countries of origin. 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).

[23] In our baseline specification, column 5 of Table 4, the beta coefficient is 0.0121. Note how without family fixed effects, the immigrant performance index is three times larger than in the specification with family fixed effects, probably capturing part of the selection into schools. Indeed, the inclusion of this variable reduces the selection effect, as shown by comparing the coefficients of our main variable across specifications: the first column of Table 4 had a negative and significant coefficient of immigrant exposure, while the equivalent coefficient in the first column of Table 10 is positive and insignificant.

To study the potential impact of disciplinary behavior of immigrants, Table 11 repeats the same exercise by constructing an immigrant performance index based on the average disciplinary behavior of the immigrants by country of origin.[24] Exposure to better-behaved immigrants has a positive and significant effect on academic outcomes, albeit small (beta coefficient is -0.006). This channel does not affect the direct impact of immigrant exposure: the beta coefficient of this variable remains similar (0.011) to the baseline specification (0.0121).[25]

When we split the sample by SES (Tables 12A and 12B), we find some interesting results in comparison to Tables 6 and 7. The immigrant performance index based on math scores has very similar beta coefficient in both subsamples, suggesting that the absolute performance of the immigrants has a consistent positive effect on all US-born students, while our main variable (immigrant exposure) behaves exactly as in Tables 6 and 7: positive and significant for FRPL eligible and Black students, null and insignificant for White and FRPL ineligible students.[26]

Overall, these results suggest that the presence of immigrants with higher academic performance correlates with better scores of US-born students. Even after controlling for absolute performance of immigrant students, a higher fraction of immigrants is still associated with higher achievement of US-born students, concentrated among those students exposed to immigrants who, in relative terms, perform better than they do. Remarkably, immigrant students do not negatively affect US-born students, even when their academic achievement is relatively lower.

5. Alternative Interpretations

It is possible that when many immigrant students attend the same school, they are "segregated" in special classes, for example because these students take remedial English classes while US-born students attend separate classes with potentially better targeted resources. If this is true, our results may simply reflect lower availability of resources in schools with fewer immigrants or may coincide artificially with less exposure of US-born students to immigrants. To investigate if this is the case, we make use of aggregate school-level measures of the classroom distribution of students,

---

[24] Lazear (2001) presents a disruption model of education in which individual disruption negatively affects the production of education.

[25] The distribution of the country of origin disciplinary behavior (Figure A6, Panel B in the Appendix) confirm large 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).

[26] When we use the immigrant performance index based on disciplinary incidents, the results are very similar (Table A13 in the Appendix).

provided by the Florida Department of Education. We calculate a 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 US-born students in each classroom in the school, USB_{syg} is the number of US-born students in the school. We first present in Figure A7 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 Figure A8, we plot the histograms of foreign-born exposure for those schools with segregation level above and below the median. In Table A14, we present a regression analysis in which we explore whether segregation is potentially a threat for our interpretation of the results. In column 1, we weight the cumulative exposure coefficient by segregation and re-estimate the Column 5 model of Table 4 for math scores. The results show a beta coefficient slightly smaller 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 beta coefficient is always higher in the sub-samples of schools with lower segregation. Because these data are aggregate-level measures only, and not at the individual level, we are unable to study other specific classroom effects, but the analyses we are able to conduct with the available data invariably indicate that immigrant segregation within schools is not responsible for our results.

A second explanation is that presence of immigrants increases the overall school diversity in cohorts with higher immigrant shares and our regression is indirectly capturing a positive impact of diversity on learning. To study this hypothesis we create a measure of diversity based on race following the literature on political economy (e.g., Alesina et al., 2003), as one minus the Herfindahl index of students by race. We calculate this measure for the entire school population and compute its cumulative counterpart, following equation (1). Table A15, column 1 presents the results adding this diversity measure which is positive but not statistically significant. The coefficient of our measure of immigrant exposure is unaffected.

Another possibility is that a higher share of immigrants is proxying for higher diversity of immigrants, and that it is the latter to be relevant for the academic achievement of the US-born students. To address this possibility, 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. Columns 2 and 3 report regressions results where we include these controls in our baseline. Both measures of diversity are positive and insignificant, and the coefficient of our main explanatory variable is unaffected, ruling out the interpretation that immigrant exposure has an impact on academic achievement through increased diversity.

6. Conclusions

We study the effect of exposure to immigrants on educational outcomes of US-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 school selection problems, comparing the test scores in math and reading of siblings who experience different school-cohort-specific immigrant concentrations, holding the heterogeneity of the families' life cycles fixed.

Our main result points to a strong selection of US-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 US-born and higher-SES students consistently with the *white flight* literature: White native students are more likely to flee schools that attract a large fraction of immigrants.

Our identification strategy provides new results about the effects of immigrants on the educational outcomes of US-born students: once selection is accounted for with family fixed effects, the correlation between cumulative immigrant exposure and academic achievement of US-born students is positive and significant. Moving from the 10^{th} to the 90^{th} percentile in the distribution of cumulative exposure (1% and 13%, respectively) increases the score in mathematics and reading by 2.8% and 1.7% 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, suggesting that immigrant students do not negatively affect US-born students, even when immigrants' academic achievement is lower than the US-born schoolmates. Even after controlling for absolute performance of immigrant students, a higher fraction of immigrants is still associated with higher achievement of US-born students, concentrated among those students exposed to immigrants who, in relative terms, perform better than they do. This finding suggests that our results may be driven by the relative differences in performance and behavior of immigrants and their US-born schoolmates. Overall, the presence of immigrants benefits disadvantaged US-born students and does not negatively affect affluent US-born students.

References

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

Ashraf, Q., O. Galor, and M. Klemp (2015), "Heterogeneity and Productivity," Working Papers 2015-4, Brown University, Department of Economics.

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.

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.

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

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,

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.

Li, M. (2009), "Is There "White Flight" into Private Schools? New Evidence from High School and Beyond," *Economics of Education Review* 28: 382–392.

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.

Morales, N. C., Do Refugee Students Affect the Academic Achievement of Peers? Evidence from a Large Urban School District (November 16, 2020). Available at SSRN: https://ssrn.com/abstract=3731794

Miller, D., N. Shenhav and M. Grosz (2019), "Selection into Identification in Fixed Effects Models with Application to Head Start," working paper 26174, National Bureau of Economic Research.

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.

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

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

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.

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

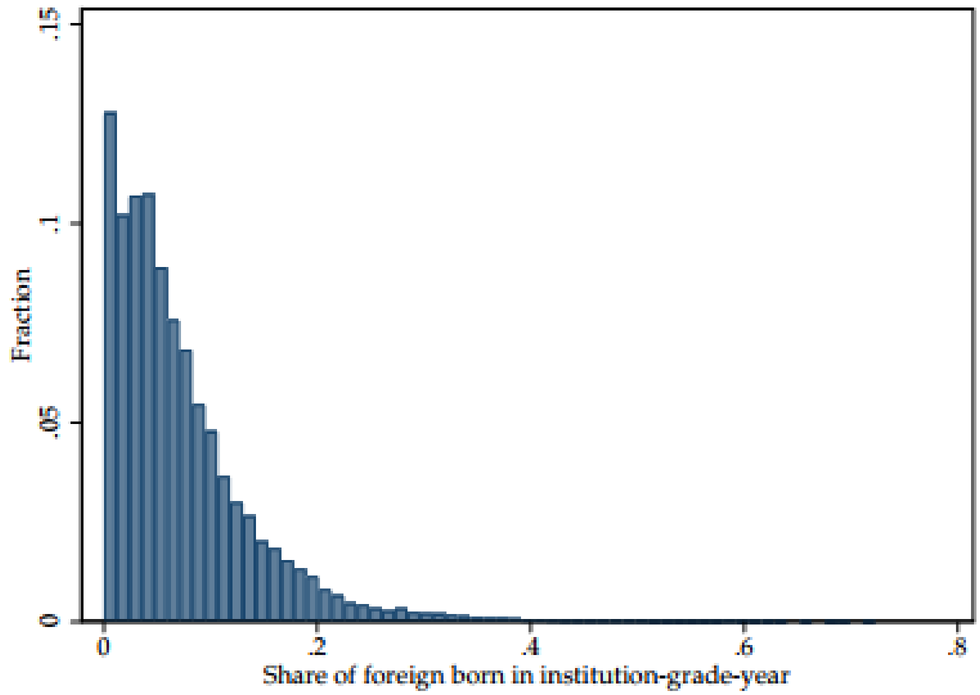

Figure 1: Distribution of foreign-born students as a share (a measure in the interval [0,1]) of the total number of students across institution-grade-year cells. The y-axis refers to the fraction of observations corresponding to values of the x-axis. The reference sample of US-born students is an unbalanced longitudinal sample of US-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.

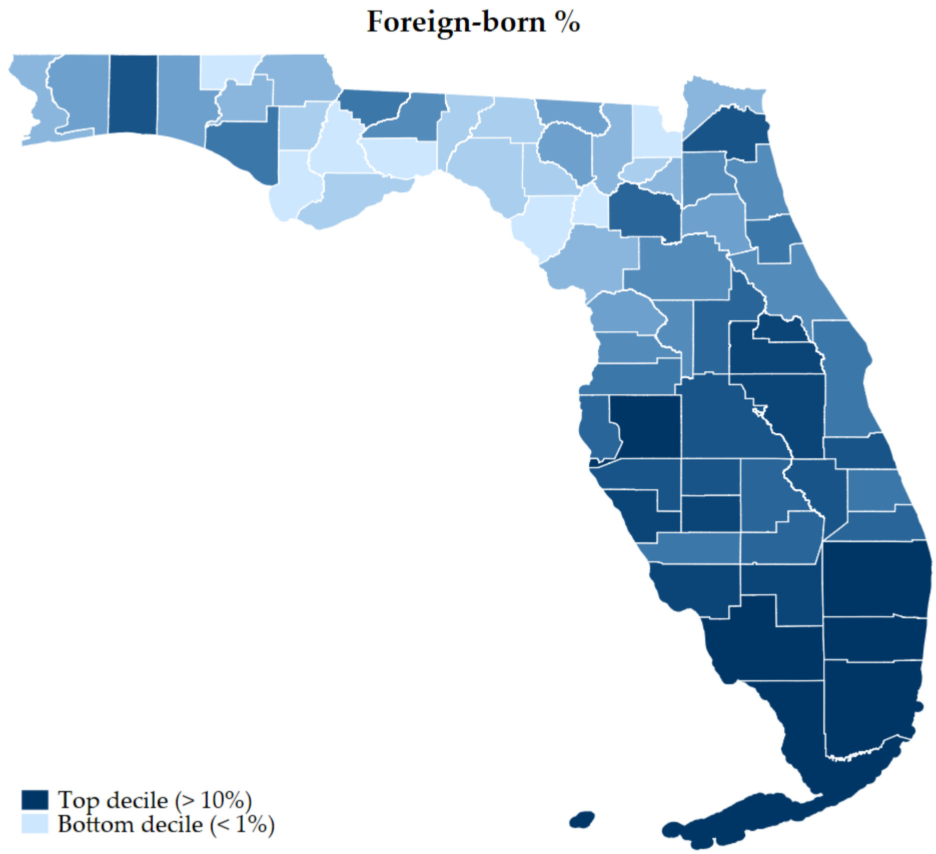

Figure 2: Concentration of foreign-born students across school districts in Florida. For each district we computed the percentage of foreign-born students over the total population of students across all years in the sample. The distribution across districts has been split in deciles and each gradation 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 US-born students is an unbalanced longitudinal sample of US-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.

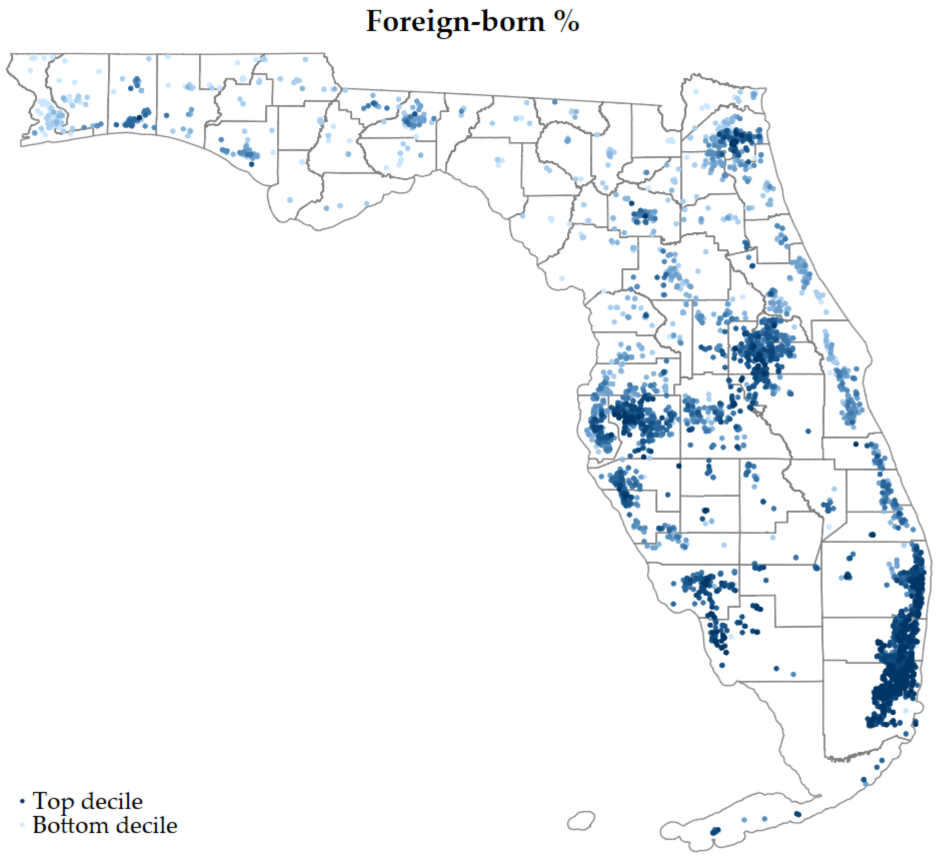

Figure 3A: Each dot in the map corresponds to one educational institution (a school). For each school we computed the percentage of foreign-born students over the total population of students across all years in the sample. The distribution across schools has been split in deciles and each gradation 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 US-born students is an unbalanced longitudinal sample of US-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.

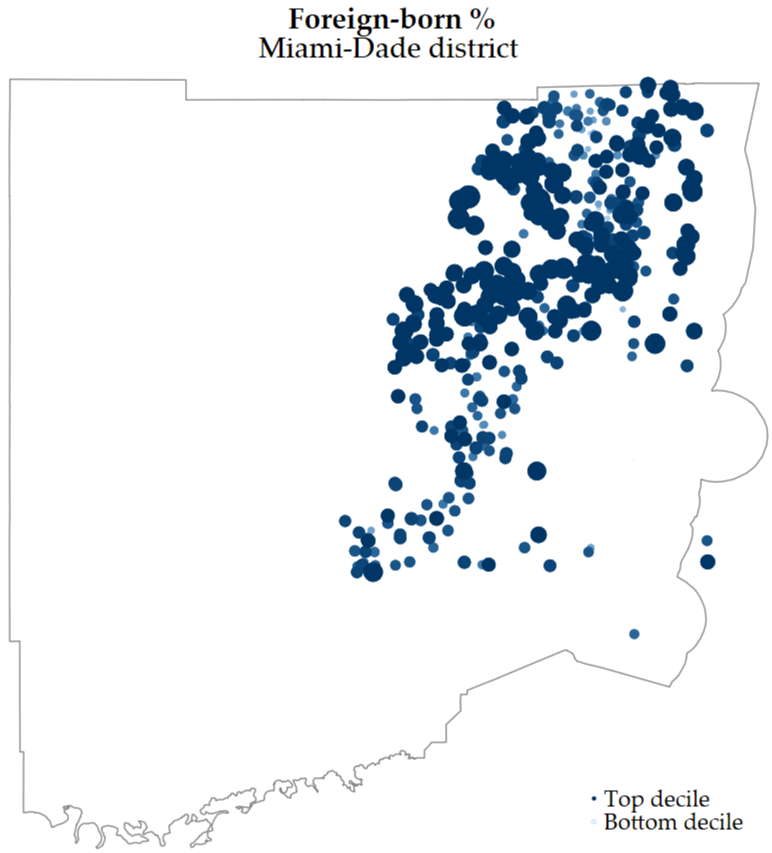

Figure 3B: Each dot corresponds to an educational institution in the Miami-Dade school district. The meaning of the color is the same as in Figure 3A, meaning that lighter colors correspond to lower deciles in the distribution of foreign-born students concentration in the whole state of Florida. The size of the dots corresponds to the size of the student body. The reference sample of US-born students is an unbalanced longitudinal sample of US-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.

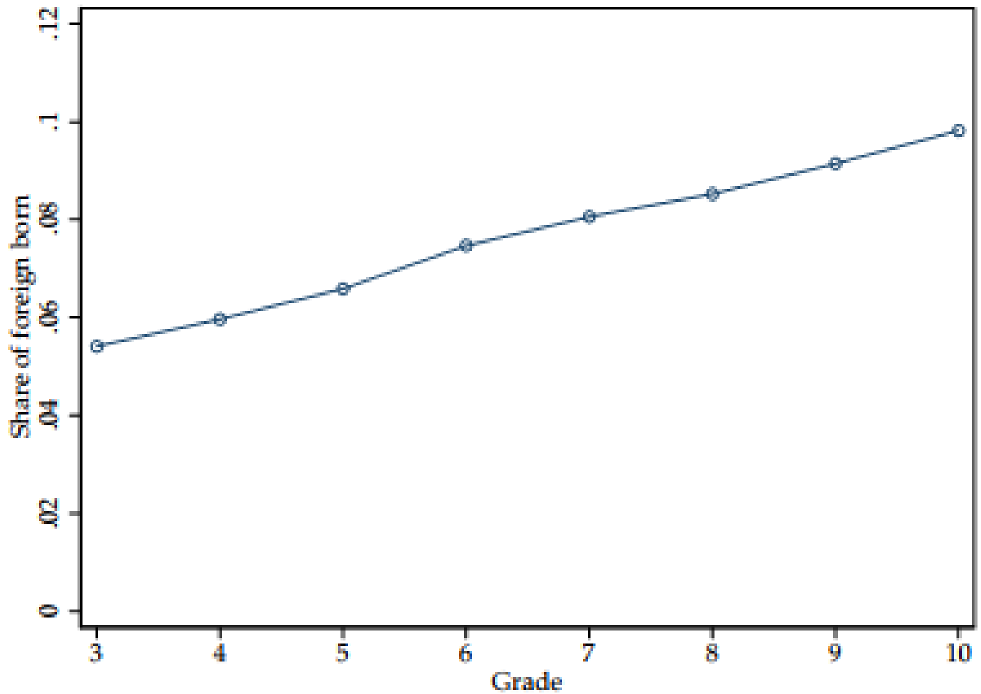

Figure 4A: Using observations across the entire time span available in the data (2002-2011), we compute the average share of foreign-born classmates for US-born English-speaking students, for each grade from 3 to 10. The reference sample of US-born students is an unbalanced longitudinal sample of US-born 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.

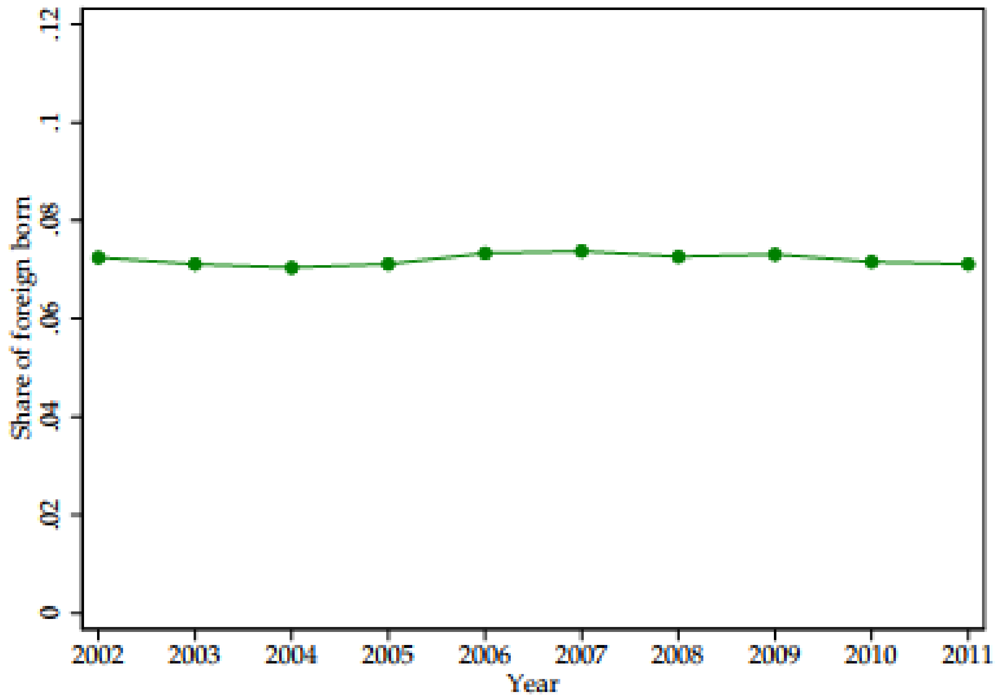

Figure 4B: Using observations across the entire time span available in the data (2002-2011), we compute the average share of foreign-born classmates for US-born English-speaking students, for each year from 2002 to 2011. The reference sample of US-born students is an unbalanced longitudinal sample of US-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.

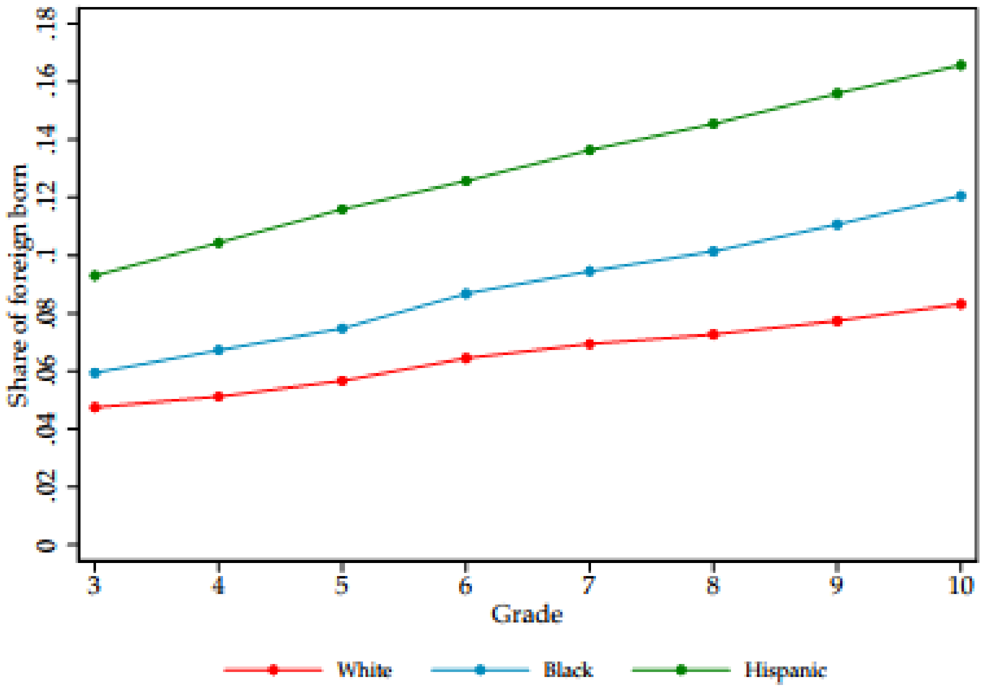

Figure 5: Using observations across the entire time span available in the data (2002-2011), we compute the average share of foreign-born classmates for three major racial/ethnic groups of US-born English-speaking students, for each grade from 3 to 10 The red line shows average exposures to foreign-born students for White US-born students, the blue line shows an analogous figure for Black US-born students, and the green line does exactly the same for Hispanic US-born students. The reference sample of US-born students is an unbalanced longitudinal sample of US-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.

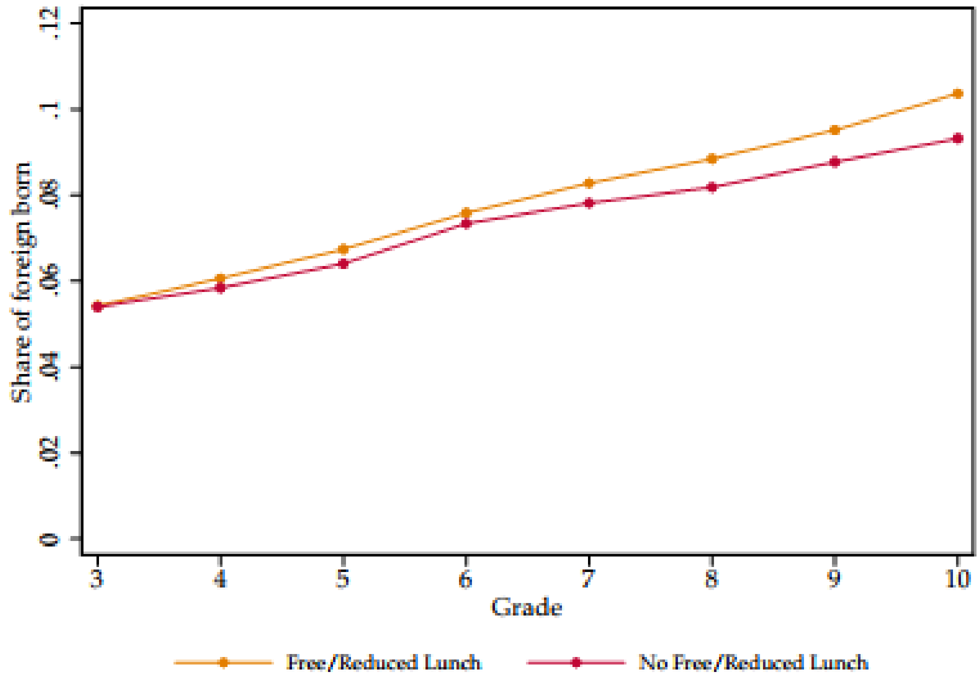

Figure 6: Using observations across the entire time span available in the data (2002-2011), we compute the average share of foreign-born classmates for two groups of US-born English-speaking students (namely, those who are eligible for free or reduced-price lunch, and those who are not), for each grade from 3 to 10. The red line shows average exposures to foreign-born students for US-born students who are not eligible for free or reduced-price lunch, while the orange line shows the same average exposure for eligible US-born students. The reference sample of US-born students is an unbalanced longitudinal sample of US-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.

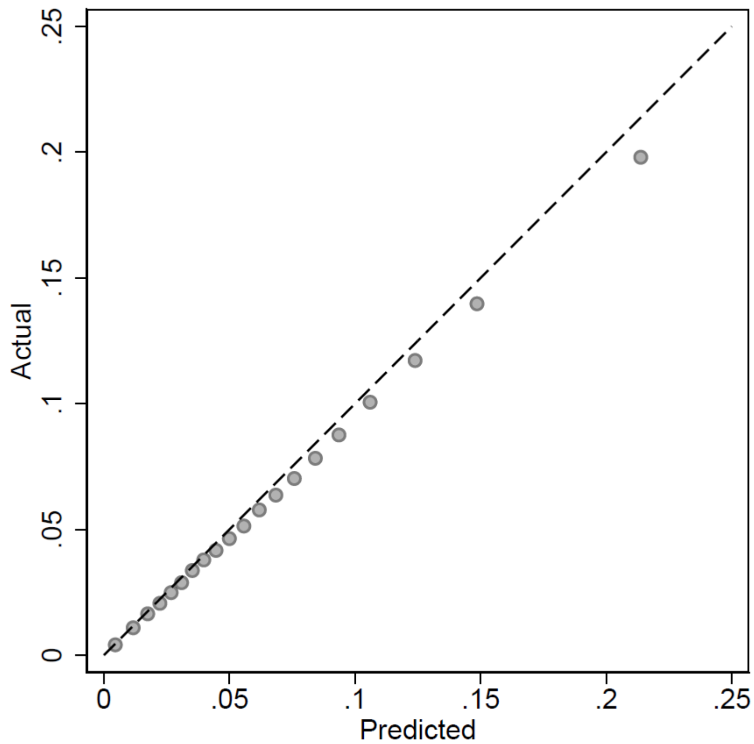

Figure 7A: This figure is a binned scatter plot that shows the raw correlation between the predicted cumulative exposure to foreign-born students and the actual cumulative exposure. Please refer to the text for details on the construction of the predicted cumulative exposure. The dashed line represents the 45-degree locus, along which the two variables are identical. The reference sample of US-born students is an unbalanced longitudinal sample of US-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).

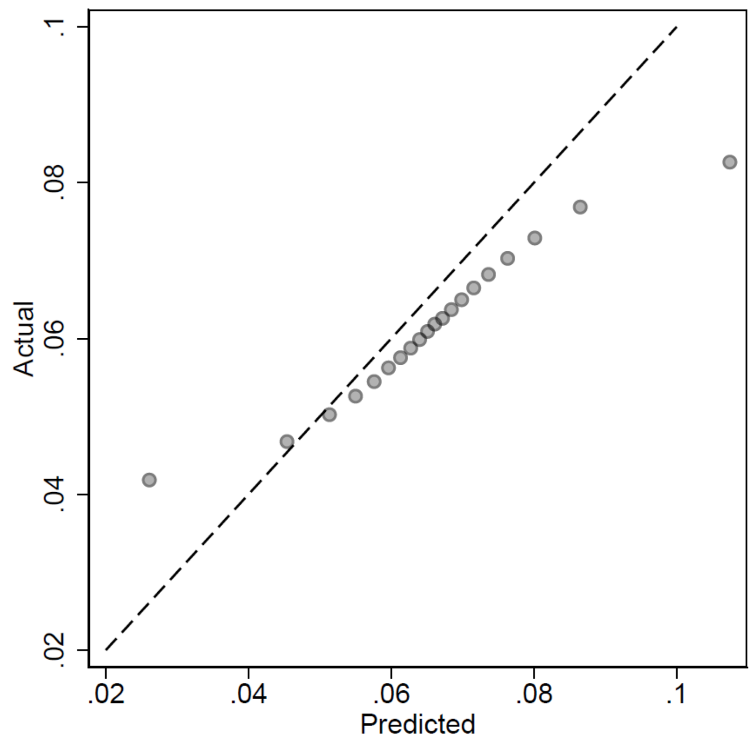

Figure 7B: This figure is a binned scatter plot that shows the correlation between the predicted cumulative exposure to foreign-born students and the actual cumulative exposure, conditional on family by initial school Fixed Effects. Please refer to the text for details on the construction of the predicted cumulative exposure. The dashed line represents the 45-degree locus, along which the two variables are identical. The reference sample of US-born students is an unbalanced longitudinal sample of US-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	Observations	Mean	Std. Dev.
Outcomes:			
Math Standardized Score	1,347,287	0.050	0.993
Reading Standardized Score	1,450,139	0.034	0.992
Incidents (ever involved in)	1,450,139	0.169	0.375
Explanatory variable of interest:			
Foreign-born Exposure	1,347,287	0.060	0.052
Individual or family characteristics:			
Female (Indicator)	1,347,287	0.498	0.500
Age in Months	1,347,287	135.5	23.2
Special Education (Indicator)	1,347,287	0.147	0.354
Birth Order	1,347,287	2.199	1.170
White (Indicator)	1,347,287	0.603	0.489
Black (Indicator)	1,347,287	0.297	0.457
Hispanic (Indicator)	1,347,287	0.052	0.223
Asian (Indicator)	1,347,287	0.007	0.082
Other (Indicator)	1,347,287	0.042	0.200
Free/Reduced-Price Lunch (Indicator)	1,347,287	0.546	0.498
Limited English Proficiency (Indicator)	1,347,287	0.002	0.043
Mother High School Drop-out (Indicator)	1,344,542	0.200	0.400
Mother High School Graduate (Indicator)	1,344,542	0.367	0.482
Mother Some College (Indicator)	1,344,542	0.239	0.426
Mother 4-year College or more (Indicator)	1,344,542	0.194	0.396

Table 1: Summary statistics. Each variable is measured on observations such that the score in mathematics is non missing; except the reading score and the incident variables, which are measured whenever available. 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. The acronyms FRL and LEP indicate ‘Free/Reduced-price Lunch’ and ‘Limited English Proficiency’, respectively. All statistics are computed on an unbalanced longitudinal sample of US-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.

	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 US-born students. White/Hispanic/Black majority indicates that only school-specific cohorts with more than 50% US-born of that specific race/ethnicity are selected. The reference sample of US-born students is an unbalanced longitudinal sample of US-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 US born students. White/Hispanic/Black majority indicates that only school-specific cohorts with more than 50% US-born of that specific race/ethnicity are selected. All statistics are computed on an unbalanced longitudinal sample of US-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.

	Math standardized score (3rd-10th grade)				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	-0.123**	0.019	0.077*	0.293***	0.229***
	(0.053)	(0.042)	(0.040)	(0.054)	(0.074)
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
Observations	1,347,287	1,347,287	1,344,542	1,347,287	1,347,287
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
Mean RHS	0.0604	0.0604	0.0604	0.0604	0.0604
SD RHS	0.0523	0.0523	0.0523	0.0523	0.0523
Standardized Coefficient	-0.00648	0.00102	0.00406	0.0154	0.0121

Table 4: This table shows the estimates of a linear regression of test scores in mathematics 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 US-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. *** p<0.01, ** p<0.05, * p<0.1.

	Reading standardized score (3rd-10th grade)				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	-0.194***	-0.026	0.040	0.176***	0.110*
	(0.049)	(0.039)	(0.037)	(0.048)	(0.064)
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
Observations	1,450,139	1,450,139	1,447,279	1,450,139	1,450,139
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
Mean RHS	0.0614	0.0614	0.0614	0.0614	0.0614
SD RHS	0.0528	0.0528	0.0528	0.0528	0.0528
Standardized Coefficient	-0.0103	-0.00138	0.00214	0.00934	0.00583

Table 5: This table shows the estimates of a linear regression of test scores in reading 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 US-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. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

	Math standardized score (3rd-10th grade)				
	Sample restriction: Race = 'White'				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	-0.610***	-0.395***	-0.261***	0.213***	0.128
	(0.064)	(0.061)	(0.058)	(0.075)	(0.107)
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
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
Mean RHS	0.0531	0.0531	0.0531	0.0531	0.0531
SD RHS	0.0470	0.0470	0.0470	0.0470	0.0470
Standardized Coefficient	-0.0314	-0.0204	-0.0135	0.0110	0.00661

Table 6A: This table shows the estimates of a linear regression of test scores in mathematics 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 US-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. *** p<0.01, ** p<0.05, * p<0.1.

	Math standardized score (3rd-10th grade)				
	Sample restriction: Race = 'Black'				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	0.517***	0.500***	0.481***	0.450***	0.402***
	(0.067)	(0.066)	(0.065)	(0.097)	(0.137)
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
Observations	399,586	399,586	398,269	399,586	399,586
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
Mean RHS	0.0663	0.0663	0.0664	0.0663	0.0663
SD RHS	0.0522	0.0522	0.0522	0.0522	0.0522
Standardized Coefficient	0.0283	0.0274	0.0264	0.0246	0.0220

Table 6B: This table shows the estimates of a linear regression of test scores in mathematics 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 US-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. 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. *** p<0.01, ** p<0.05, * p<0.1.

	Math standardized score (3rd-10th grade) *Sample restriction: Lunch Status = No Free/Reduced-price*				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	-0.460***	-0.424***	-0.296***	-0.002	-0.034
	(0.067)	(0.065)	(0.061)	(0.080)	(0.113)
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
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
Mean RHS	0.0601	0.0601	0.0601	0.0601	0.0601
SD RHS	0.0520	0.0520	0.0520	0.0520	0.0520
Standardized Coefficient	-0.0276	-0.0254	-0.0178	-0.0001	-0.00202

Table 7A: This table shows the estimates of a linear regression of test scores in mathematics 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 US-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. *** p<0.01, ** p<0.05, * p<0.1.

	Math standardized score (3rd-10th grade)				
	Sample restriction: Lunch Status = Free/Reduced-price				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	0.368***	0.283***	0.301***	0.452***	0.399***
	(0.053)	(0.050)	(0.049)	(0.074)	(0.102)
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
Observations	735,589	735,589	733,624	735,589	735,589
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
Mean RHS	0.0607	0.0607	0.0607	0.0607	0.0607
SD RHS	0.0525	0.0525	0.0525	0.0525	0.0525
Standardized Coefficient	0.0203	0.0156	0.0166	0.0250	0.0220

Table 7B: This table shows the estimates of a linear regression of test scores in mathematics 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 US-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. *** p<0.01, ** p<0.05, * p<0.1.

	Sample restriction: siblings who go to the same initial school Math standardized score (3rd-10th grade)		
	(1) *IV*	(3) *Red. Form*	(5) *OLS*
Foreign-born Exposure	0.319**		0.338***
	(0.155)		(0.068)
Foreign-born Exposure (Predicted)		0.139**	
		(0.067)	
Individual controls	X	X	X
Year x Grade FE	X	X	X
Family x Initial School FE	X	X	X
Observations	821,892	821,892	821,892
R-squared	-	0.668	0.668
Dependent Variable (mean)	0.192	0.192	0.192
Dependent Variable (sd)	0.954	0.954	0.954
RHS (mean)	0.062	0.066	0.062
RHS (sd)	0.055	0.052	0.055
Standardized coefficient	0.018	0.008	0.019

Table 8: This table shows results on the 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 US-born students observed in grades from 3^{rd} to 10^{th}, 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, 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. Year x grade FEs are indicators for each unique year-grade combination. Family x Initial school FEs are indicators for each unique family-initial school combination. Robust standard errors in parenthesis clustered by school-cohort. *** p<0.01, ** p<0.05, * p<0.1.

	Math Score (Standardized)		Reading Score (Standardized)		Incidents (indicator)	
	Obs.	Mean	Obs.	Mean	Obs.	Mean
sample: Whole sample						
US-born speaking English	1,347,287	0.050	1,450,139	0.034	1,450,139	0.169
Immigrants who go to school with those above	948,590	0.006	1,025,267	-0.071	948,590	0.119
sample: White US born students						
US-born speaking English	811,790	0.305	873,281	0.288	873,281	0.105
Immigrants who go to school with those above	788,626	0.093	861,388	0.026	862,803	0.110
sample: Black US born students						
US-born speaking English	399,586	-0.495	430,975	-0.511	430,975	0.310
Immigrants who go to school with those above	763,358	-0.180	837,098	-0.275	838,408	0.142
sample: No-Free/Reduced price lunch US born						
US-born speaking English	611,698	0.475	658,656	0.459	658,656	0.074
Immigrants who go to school with those above	817,046	0.170	892,540	0.101	893,879	0.104
sample: Free/Reduced price lunch US born						
US-born speaking English	735,589	-0.303	791,483	-0.319	791,483	0.247
Immigrants who go to school with those above	899,632	-0.137	976,310	-0.220	977,931	0.131

Table 9: This table shows descriptive statistics of test scores and incident rates across different subset of students. In particular, it shows the mean of each variable for the sample of US-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 US-born students and then for four different subsets, based on reported race and free lunch eligibility.

	Math standardized score (3rd-10th grade)				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	0.007	0.085**	0.122***	0.294***	0.216***
	(0.052)	(0.042)	(0.040)	(0.055)	(0.077)
Immigrant performance index (Math score)	0.308***	0.152***	0.106***	0.036***	0.040***
	(0.009)	(0.007)	(0.007)	(0.006)	(0.008)
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
Observations	1,279,001	1,279,001	1,276,539	1,279,001	1,279,001
R-squared	0.305	0.360	0.381	0.687	0.777
Mean LHS	0.0579	0.0579	0.0585	0.0579	0.0579
SD LHS	0.993	0.993	0.993	0.993	0.993
Mean RHS	0.0633	0.0633	0.0633	0.0633	0.0633
SD RHS	0.0520	0.0520	0.0520	0.0520	0.0520
Standardized Coefficient	0.000379	0.00445	0.00640	0.0154	0.0113

Table 10: This table shows the estimates of a linear regression of test scores in mathematics 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, a cumulative index of foreign-born peers' math performance (computed as the average performance index across previous school-specific cohorts including the current grade), and several controls. A school-cohort index of foreign-born performance is computed as a weighted average of country-specific mean math test scores, weighted by the share of students from a given country, in a given school-specific cohort. All regressions are run on an unbalanced longitudinal sample of US-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. *** p<0.01, ** p<0.05, * p<0.1.

	Math standardized score (3rd-10th grade)				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	-0.100*	0.032	0.086**	0.286***	0.209***
	(0.052)	(0.042)	(0.040)	(0.055)	(0.077)
Immigrant performance index (Behavior)	-1.919***	-0.864***	-0.588***	-0.182***	-0.222***
	(0.066)	(0.050)	(0.048)	(0.044)	(0.059)
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
Observations	1,279,001	1,279,001	1,276,539	1,279,001	1,279,001
R-squared	0.304	0.360	0.380	0.687	0.777
Mean LHS	0.0579	0.0579	0.0585	0.0579	0.0579
SD LHS	0.993	0.993	0.993	0.993	0.993
Mean RHS	0.0633	0.0633	0.0633	0.0633	0.0633
SD RHS	0.0520	0.0520	0.0520	0.0520	0.0520
Standardized Coefficient	-0.00524	0.00170	0.00448	0.0150	0.0110

Table 11: This table shows the estimates of a linear regression of test scores in mathematics 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, a cumulative index of foreign-born peers' behavioral performance (computed as the average performance index across previous school-specific cohorts including the current grade), and several controls. A school-cohort index of foreign-born behavioral performance is 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 regressions are run on an unbalanced longitudinal sample of US-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. *** p<0.01, ** p<0.05, * p<0.1.

	Math standardized score (3rd-10th grade) *Sample restriction: Lunch Status = No Free/Reduced-price*				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	-0.374***	-0.351***	-0.243***	-0.017	-0.061
	(0.067)	(0.065)	(0.062)	(0.082)	(0.117)
Immigrant performance index (Math score)	0.192***	0.160***	0.106***	0.028***	0.035***
	(0.010)	(0.010)	(0.010)	(0.009)	(0.013)
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
Observations	587,588	587,588	586,877	587,588	587,588
R-squared	0.219	0.236	0.271	0.675	0.769
Mean LHS	0.480	0.480	0.480	0.480	0.480
SD LHS	0.867	0.867	0.867	0.867	0.867
Mean RHS	0.0624	0.0624	0.0624	0.0624	0.0624
SD RHS	0.0518	0.0518	0.0518	0.0518	0.0518
Standardized Coefficient	-0.0224	-0.0209	-0.0145	-0.00100	-0.00362

Table 12A: This table shows the estimates of a linear regression of test scores in mathematics 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, a cumulative index of foreign-born peers' math performance (computed as the average performance index across previous school-specific cohorts including the current grade), and several controls. A school-cohort index of foreign-born performance is computed as a weighted average of country-specific mean math test scores, weighted by the share of students from a given country, in a given school-specific cohort. All regressions are run on an unbalanced longitudinal sample of US-born students not eligible for free or reduced-price lunch, 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. 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. *** p<0.01, ** p<0.05, * p<0.1.

	Math standardized score (3rd-10th grade) *Sample restriction: Lunch Status = Free/Reduced-price*				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	0.451***	0.340***	0.342***	0.452***	0.386***
	(0.053)	(0.050)	(0.049)	(0.076)	(0.107)
Immigrant performance index (Math score)	0.191***	0.126***	0.096***	0.038***	0.040***
	(0.009)	(0.008)	(0.008)	(0.008)	(0.011)
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
Observations	691,413	691,413	689,662	691,413	691,413
R-squared	0.251	0.280	0.294	0.627	0.739
Mean LHS	-0.301	-0.301	-0.301	-0.301	-0.301
SD LHS	0.952	0.952	0.952	0.952	0.952
Mean RHS	0.0641	0.0641	0.0642	0.0641	0.0641
SD RHS	0.0522	0.0522	0.0522	0.0522	0.0522
Standardized Coefficient	0.0247	0.0186	0.0188	0.0248	0.0212

Table 12B: This table shows the estimates of a linear regression of test scores in mathematics 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, a cumulative index of foreign-born peers' math performance (computed as the average performance index across previous school-specific cohorts including the current grade), and several controls. A school-cohort index of foreign-born performance is computed as a weighted average of country-specific mean math test scores, weighted by the share of students from a given country, in a given school-specific cohort. All regressions are run on an unbalanced longitudinal sample of US-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. *** p<0.01, ** p<0.05, * p<0.1.

Appendix

Tables

	All US-Born						*English-speaking US-born*					
	All observations			2+ siblings per year			All observations			2+ siblings per year		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variable	*Obs.*	*Mean*	*Std. Dev.*	*Obs.*	*Mean*	*Std. Dev.*	*Obs.*	*Mean*	*Std. Dev.*	*Obs.*	*Mean*	*Std. Dev.*
Outcomes:												
Math Standardized Score	7,490,949	0.044	0.971	1,662,403	0.043	0.986	5,924,346	0.071	0.969	1,347,287	0.050	0.993
Reading Standardized Score	8,010,198	0.052	0.968	1,789,450	0.016	0.983	6,341,333	0.096	0.967	1,450,139	0.034	0.992
Incidents (ever involved in)	8,010,198	0.143	0.350	1,789,450	0.162	0.368	6,341,333	0.149	0.356	1,450,139	0.169	0.375
Explanatory variable of interest:												
Foreign-born Exposure	7,490,949	0.079	0.070	1,662,403	0.074	0.067	5,924,346	0.065	0.057	1,347,287	0.060	0.052
Individual or family characteristics:												
Female (Indicator)	7,490,949	0.495	0.500	1,662,403	0.498	0.500	5,924,346	0.495	0.500	1,347,287	0.498	0.500
Age in Months	7,490,949	131.9	23.6	1,662,403	135.4	23.2	5,924,346	132.1	23.6	1,347,287	135.5	23.2
Special Education (Indicator)	7,490,949	0.138	0.345	1,662,403	0.145	0.352	5,924,346	0.139	0.346	1,347,287	0.147	0.354
Birth Order	7,490,949	1.985	1.142	1,662,403	2.201	1.179	5,924,346	1.973	1.123	1,347,287	2.199	1.170
White Student (Indicator)	7,490,949	0.493	0.500	1,662,403	0.509	0.500	5,924,346	0.601	0.490	1,347,287	0.603	0.489
Black (Indicator)	7,490,949	0.225	0.418	1,662,403	0.271	0.444	5,924,346	0.255	0.436	1,347,287	0.297	0.457
Hispanic (Indicator)	7,490,949	0.217	0.412	1,662,403	0.165	0.371	5,924,346	0.082	0.274	1,347,287	0.052	0.223
Asian (Indicator)	7,490,949	0.020	0.141	1,662,403	0.014	0.116	5,924,346	0.013	0.112	1,347,287	0.007	0.082
Other (Indicator)	7,490,949	0.045	0.207	1,662,403	0.041	0.198	5,924,346	0.049	0.217	1,347,287	0.042	0.200
Free/Reduced-Price Lunch (Indicator)	7,490,186	0.536	0.499	1,662,403	0.579	0.494	5,923,759	0.486	0.500	1,347,287	0.546	0.498
Limited English Proficiency (Indicator)	7,490,949	0.038	0.190	1,662,403	0.019	0.136	5,924,346	0.004	0.066	1,347,287	0.002	0.043
Mother High School DO (Indicator)	5,219,361	0.224	0.417	1,658,296	0.219	0.414	4,164,506	0.194	0.395	1,344,542	0.200	0.400
Mother High School Graduate (Indicator)	5,219,361	0.376	0.484	1,658,296	0.365	0.481	4,164,506	0.381	0.486	1,344,542	0.367	0.482
Mother Some College (Indicator)	5,219,361	0.234	0.423	1,658,296	0.232	0.422	4,164,506	0.249	0.432	1,344,542	0.239	0.426
Mother 4-year College or more (Indicator)	5,219,361	0.166	0.372	1,658,296	0.185	0.388	4,164,506	0.176	0.381	1,344,542	0.194	0.396

Table A1.A: Summary statistics of US born students. All statistics are computed on an unbalanced sample of students born between 1994 and 2002, observed in any grade between 3 and 10. Each variable is measured on observations such that the score in mathematics is non-missing; except the reading score and the incident variables, which are measured whenever available. Columns 1-6 show summary statistics computed on the entire sample of observations of US-born students in (columns 1-3), and on the restricted sample of observations such that at least two siblings are observed in a given year in (4-6). In columns 7-12 we do the same exercise for US-born students speaking English. Columns (10-12) contain our main sample and it is identical to Table 1 in the text. 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.

	Foreign-born peers of all US-Born						*Foreign-born peers of English-speaking US-born*					
	All observations			**US-fam. 2+ siblings per year**			**All observations**			**US-fam. 2+ siblings per year**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variable	*Obs.*	*Mean*	*Std. Dev.*	*Obs.*	*Mean*	*Std. Dev.*	*Obs.*	*Mean*	*Std. Dev.*	*Obs.*	*Mean*	*Std. Dev.*
Outcomes:												
Math Standardized Score	875,990	-0.097	1.109	854,867	-0.093	1.107	875,585	-0.097	1.109	830,857	-0.087	1.106
Reading Standardized Score	946,117	-0.206	1.142	924,771	-0.202	1.141	945,642	-0.206	1.142	900,324	-0.195	1.138
Incidents (ever involved in)	946,117	0.127	0.333	924,771	0.128	0.335	945,642	0.126	0.333	900,324	0.130	0.337
Explanatory variable of interest:												
Foreign-born Exposure	875,990	0.176	0.107	854,867	0.176	0.107	875,585	0.176	0.107	830,857	0.173	0.106
Individual or family characteristics:												
Female (Indicator)	875,990	0.490	0.500	854,867	0.490	0.500	875,585	0.490	0.500	830,857	0.490	0.500
Age in Months	875,990	137.8	25.8	854,867	138.4	25.6	875,585	137.8	25.8	830,857	139.0	25.6
Special Education (Indicator)	875,990	0.087	0.282	854,867	0.087	0.282	875,585	0.087	0.282	830,857	0.087	0.282
Birth Order	875,990	2.160	1.355	854,867	2.154	1.359	875,585	2.160	1.355	830,857	2.157	1.360
White (Indicator)	875,990	0.131	0.337	854,867	0.130	0.336	875,585	0.131	0.337	830,857	0.133	0.339
Black (Indicator)	875,990	0.166	0.372	854,867	0.166	0.372	875,585	0.166	0.372	830,857	0.169	0.374
Hispanic (Indicator)	875,990	0.614	0.487	854,867	0.614	0.487	875,585	0.614	0.487	830,857	0.607	0.488
Asian (Indicator)	875,990	0.068	0.252	854,867	0.068	0.252	875,585	0.068	0.252	830,857	0.070	0.255
Other (Indicator)	875,990	0.022	0.145	854,867	0.021	0.145	875,585	0.022	0.145	830,857	0.022	0.146
Free/Reduced-Price Lunch (Indicator)	875,829	0.682	0.466	854,708	0.682	0.466	875,424	0.682	0.466	830,704	0.677	0.467
Limited English Proficiency (Indicator)	875,990	0.321	0.467	854,867	0.318	0.466	875,585	0.321	0.467	830,857	0.315	0.464

Table A1.B: Summary statistics of immigrant students. The summary statistics displayed are computed on the sample of foreign-born peers going to school with different groups of US-born students. Columns 1-3 shows summary statistics computed on the sample of foreign-born peers of all US-born students. Columns 4-6 shows the same statistics for the restricted sample of observations of foreign-born peers going to school with US-born students in families such that at least two siblings are observed in a given year. In columns 7-12 we repeat the same exercise after restricting to foreign-born peers going to school with US-born students speaking English at home. 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. Each variable is measured on observations such that the score in mathematics is non-missing; except the reading score variable, which is measured for observations such that the reading score is non missing.

Panel A: Enrollment in Public School						
	US born students		1st generation		2nd generation	
	Obs.	Mean	Obs.	Mean	Obs.	Mean
Census 2000 (5%)						
Kindergarten	6,415	82.29%	646	84.83%	2,582	81.14%
Grade 1 to 4	26,500	86.69%	3,279	93.44%	9,438	86.76%
Grade 5 to 8	26,581	87.86%	4,477	93.52%	8,244	87.58%
Grade 9 to 12	21,813	90.58%	5,289	93.67%	6,576	87.61%
Overall sample	81,309	87.77%	13,691	93.15%	26,840	86.68%
Census 2010 (1%)						
Kindergarten	1,147	82.65%	91	74.73%	632	83.23%
Grade 1 to 4	4,556	85.45%	557	89.77%	2,301	88.57%
Grade 5 to 8	5,047	85.56%	855	90.64%	2,036	87.18%
Grade 9 to 12	4,726	87.85%	1,114	92.91%	1,861	88.07%
Overall sample	15,476	86.01%	2,617	90.87%	6,830	87.53%

Panel B: Family Income (USD)						
	US born students		1st generation		2nd generation	
	Obs.	Mean	Obs.	Mean	Obs.	Mean
Census 2000 (5%)						
Public school	71,364	55,838	12,648	43,526	23,264	52,842
Private school	9,945	102,409	928	86,163	3,576	106,669
Overall sample	81,309	61,534	13,576	46,441	26,840	60,014
Census 2010 (1%)						
Public school	13,311	71,906	2,372	54,343	5,978	65,630
Private school	2,165	123,921	238	115,190	852	136,119
Overall sample	15,476	79,183	2,610	59,892	6,830	74,423

Table A2: This table reports the fraction of students by grade and family income enrolled in public and private schools in Florida. The data are based on Census 2000 and 2010 and report the statistics for US-born students, first generation and second generation immigrant students. "2nd generation" is identified as having at least the mother or the father born abroad.

	Math standardized score (3rd-10th grade)				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	-0.001	0.087**	0.124***	0.321***	0.250***
	(0.055)	(0.044)	(0.042)	(0.056)	(0.077)
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
Observations	1,347,287	1,347,287	1,344,542	1,347,287	1,347,287
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
Mean RHS	0.0563	0.0563	0.0563	0.0563	0.0563
SD RHS	0.0504	0.0504	0.0504	0.0504	0.0504
Standardized Coefficient	-4.74e-05	0.00443	0.00627	0.0163	0.0127

Table A3: This table shows the estimates of a linear regression of test scores in mathematics standardized by year and grade on the cumulative exposure to foreign-born students (excluding Puerto-Rican 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 US-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. *** p<0.01, ** p<0.05, * p<0.1.

	Reading standardized score (3rd-10th grade)				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	-0.059	0.055	0.099**	0.210***	0.139**
	(0.052)	(0.041)	(0.039)	(0.050)	(0.067)
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
Observations	1,450,139	1,450,139	1,447,279	1,450,139	1,450,139
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
Mean RHS	0.0572	0.0572	0.0572	0.0572	0.0572
SD RHS	0.0508	0.0508	0.0508	0.0508	0.0508
Standardized Coefficient	-0.00304	0.00282	0.00507	0.0108	0.00710

Table A4: This table shows the estimates of a linear regression of test scores in reading standardized by year and grade on the cumulative exposure to foreign-born students (excluding Puerto-Rican 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 US-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. *** p<0.01, ** p<0.05, * p<0.1.

	Math standardized score (3rd-10th grade)				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	0.073**	0.118***	0.148***	0.245***	0.190***
	(0.030)	(0.029)	(0.029)	(0.034)	(0.046)
Math Score in 3rd Grade	0.764***	0.742***	0.732***	0.657***	0.656***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
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
Observations	1,275,020	1,275,020	1,272,415	1,275,020	1,275,020
R-squared	0.687	0.691	0.693	0.788	0.867
Mean LHS	0.0558	0.0558	0.0564	0.0558	0.0558
SD LHS	0.987	0.987	0.987	0.987	0.987
Mean RHS	0.0652	0.0652	0.0652	0.0652	0.0652
SD RHS	0.0574	0.0574	0.0574	0.0574	0.0574
Standardized Coefficient	0.00426	0.00685	0.00862	0.0142	0.0110

Table A5 This table shows estimates from models equivalent to those reported in Table 4, except that (i) the score in mathematics in 3rd grade is included as an explanatory variable; (ii) the sample is restricted to a subset of observations that exclude the 3rd grade. Robust standard errors in parenthesis clustered by school-cohort. *** p<0.01, ** p<0.05, * p<0.1.

	Math standardized score (3rd-10th grade) *Sample: US born students speaking any language at home*				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	-0.226***	0.003	0.086***	0.230***	0.161***
	(0.042)	(0.033)	(0.032)	(0.044)	(0.061)
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
Observations	1,662,403	1,662,403	1,658,296	1,662,403	1,662,403
R-squared	0.289	0.342	0.360	0.675	0.763
Mean LHS	0.0430	0.0430	0.0437	0.0430	0.0430
SD LHS	0.986	0.986	0.986	0.986	0.986
Mean RHS	0.0744	0.0744	0.0744	0.0744	0.0744
SD RHS	0.0666	0.0666	0.0666	0.0666	0.0666
Standardized Coefficient	-0.0153	0.000216	0.00578	0.0155	0.0109

Table A6: This table shows estimates from models equivalent to those reported in Table 4, except that the sample of US-born students is not restricted based on language spoken at home. This table shows the estimates of a linear regression of test scores in mathematics 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 US-born students (speaking any language at home) observed in grades from 3rd to 10th, who 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. *** p<0.01, ** p<0.05, * p<0.1.

	Reading standardized score (3rd-10th grade)				
	Sample: US born students speaking any language at home				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	-0.372***	-0.064**	0.026	0.227***	0.169***
	(0.039)	(0.032)	(0.030)	(0.039)	(0.052)
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
Observations	1,789,450	1,789,450	1,785,147	1,789,450	1,789,450
R-squared	0.292	0.341	0.361	0.661	0.746
Mean LHS	0.0158	0.0158	0.0165	0.0158	0.0158
SD LHS	0.983	0.983	0.983	0.983	0.983
Mean RHS	0.0756	0.0756	0.0756	0.0756	0.0756
SD RHS	0.0672	0.0672	0.0672	0.0672	0.0672
Standardized Coefficient	-0.0254	-0.00440	0.00177	0.0155	0.0115

Table A7: This table shows estimates from models equivalent to those reported in Table 5, except that the sample of US-born students is not restricted based on language spoken at home. This table shows the estimates of a linear regression of test scores in reading 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 US-born students (speaking any language at home) observed in grades from 3rd to 10th, who 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. *** p<0.01, ** p<0.05, * p<0.1.

	Math stdz score	Reading stdz score
	Sample: 3rd to 10th grade	
	(1)	(2)
Foreign-born Exposure	0.357***	0.209***
	(0.080)	(0.070)
Blacks' Exposure	-0.069*	-0.130***
	(0.038)	(0.032)
Hispanics' Exposure	-0.086*	-0.050
	(0.050)	(0.044)
Asians' Exposure	0.321**	0.459***
	(0.129)	(0.118)
Free-Lunch Exposure	-0.244***	-0.229***
	(0.033)	(0.030)
Limited English Proficiency Exposure	-0.079	-0.128***
	(0.053)	(0.047)
Observations	1,347,287	1,450,139
R-squared	0.769	0.752
Mean LHS	0.050	0.034
SD LHS	0.993	0.992
Mean RHS	0.060	0.061
	0.052	0.053
Standardized coefficient	0.019	0.011

Table A8: This table shows estimates from models equivalent to those reported in column 5 of Table 4 and Table 5, but adding controls for other exposures. Column (1) shows the estimates of a linear regression of test scores in mathematics 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 individual controls. Additionally, it includes controls for the cumulative exposure to Black students, the cumulative exposure to Hispanic students, the cumulative exposure to Asian students, the cumulative exposure to students enrolled in the Free Lunch program, and the cumulative exposure to students enrolled in the Limited English Proficiency program. Column (2) shows the estimates of a linear regression of test scores in reading 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 individual controls. Additionally, it includes controls for the cumulative exposure to Black students, the cumulative exposure to Hispanic students, the cumulative exposure to Asian students, the cumulative exposure to students enrolled in the Free Lunch program, and the cumulative exposure to students enrolled in the Limited English Proficiency program. All regressions are run on an unbalanced longitudinal sample of US-born students speaking English at home, observed in grades from 3rd to 10th, who 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. *** p<0.01, ** p<0.05, * p<0.1.

	Reading standardized score (3rd-10th grade)				
	Sample restriction: Race = 'White'				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	-0.759***	-0.528***	-0.378***	0.044	-0.009
	(0.062)	(0.059)	(0.056)	(0.073)	(0.099)
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
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
Mean RHS	0.0539	0.0539	0.0539	0.0539	0.0539
SD RHS	0.0473	0.0473	0.0473	0.0473	0.0473
Standardized Coefficient	-0.0385	-0.0268	-0.0191	0.00224	-0.000439

Table A9_A: This table shows the estimates of a linear regression of test scores in reading 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 US-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. *** p<0.01, ** p<0.05, * p<0.1.

	Reading standardized score (3rd-10th grade)				
	Sample restriction: Race = 'Black'				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	0.563***	0.551***	0.533***	0.371***	0.286***
	(0.059)	(0.058)	(0.058)	(0.082)	(0.110)
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
Observations	430,975	430,975	429,598	430,975	430,975
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
Mean RHS	0.0677	0.0677	0.0677	0.0677	0.0677
SD RHS	0.0528	0.0528	0.0527	0.0528	0.0528
Standardized Coefficient	0.0328	0.0322	0.0311	0.0217	0.0167

Table A9_B: This table shows the estimates of a linear regression of test scores in reading 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 US-born Black 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. 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. *** p<0.01, ** p<0.05, * p<0.1.

	Reading standardized score (3rd-10th grade)				
	Sample restriction: Lunch Status = No Free/Reduced-price				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	-0.460***	-0.424***	-0.296***	-0.002	-0.034
	(0.067)	(0.065)	(0.061)	(0.080)	(0.113)
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
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
Mean RHS	0.0601	0.0601	0.0601	0.0601	0.0601
SD RHS	0.0520	0.0520	0.0520	0.0520	0.0520
Standardized					-
Coefficient	-0.0276	-0.0254	-0.0178	-9.63e-05	0.00202

Table A10_A: This table shows the estimates of a linear regression of test scores in reading 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 US-born students ineligible 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. *** p<0.01, ** p<0.05, * p<0.1.

	Reading standardized score (3rd-10th grade) *Sample restriction: Lunch Status = Free/Reduced-price*				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	0.368***	0.283***	0.301***	0.452***	0.399***
	(0.053)	(0.050)	(0.049)	(0.074)	(0.102)
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
Observations	735,589	735,589	733,624	735,589	735,589
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
Mean RHS	0.0607	0.0607	0.0607	0.0607	0.0607
SD RHS	0.0525	0.0525	0.0525	0.0525	0.0525
Standardized Coefficient	0.0203	0.0156	0.0166	0.0250	0.0220

Table A10_B: This table shows the estimates of a linear regression of test scores in reading 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 US-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. *** p<0.01, ** p<0.05, * p<0.1.

	Math standardized score (3rd-10th grade)		
	(1)	(2)	(3)
Sample	**All**	**Same school**	**Different School**
Foreign-born Exposure	0.199**	0.243**	0.173
	(0.082)	(0.114)	(0.108)
Individual Controls	X	X	X
School x Year FE	X	X	X
Grade x Year FE	X	X	X
Family x Year FE	X	X	X
Observations	1,118,170	425,816	692,354
R-squared	0.785	0.770	0.795
Mean LHS	0.0928	0.0666	0.109
SD LHS	0.980	0.985	0.977
Mean RHS	0.0607	0.0599	0.0611
SD RHS	0.0528	0.0538	0.0523
Standardized Coefficient	0.0107	0.0132	0.00928

Table A11: This table shows estimates from a model equivalent to the one reported in column (5) of Table 4 with a different sample selection. In column (1), we include only observations of siblings in families with exactly 2 siblings in a given year. In column (2), among the observations used in column (1), we select only observations of siblings going to the same school in a given year. In column (3), among the observations used in column (1), we select only observations of siblings going to different schools in a given year. Robust standard errors in parenthesis clustered by school-cohort. *** p<0.01, ** p<0.05, * p<0.1.

	Math standardized score (3rd-10th grade)		
	(1)	(2)	(3)
	IV	*Red. Form*	*OLS*
Foreign-born Exposure	0.515***		0.221***
	(0.110)		(0.067)
Foreign-born Exposure (Predicted)		0.280***	
		(0.060)	
Grade x Year FE	X	X	X
School x Year FE	X	X	X
Family FE	X	X	X
Observations	854,191	854,191	854,191
R-squared	-	0.688	0.688
Mean LHS	0.149	0.149	0.149
SD LHS	0.974	0.974	0.974
Mean RHS	0.060	0.062	0.060
SD RHS	0.052	0.054	0.052
Standardized coefficient	0.028	0.016	0.012
First stage (coefficient)	0.545***	-	-
First stage (se)	(0.005)	-	-

Table A12: This table shows results on the instrumental variable approach using as instrument for foreign-born exposure the exposure that the student would have had if she/he had attended the same school attended by the eldest sibling in the given grade. 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. All regressions are run on an unbalanced longitudinal sample of US-born students observed in grades from 3rd to 10th, who speak English at home and have at least one sibling. We further restrict the sample by excluding households with twins, and children whose firstborn sibling is not in our sample for a given grade. Each observation is a student-year. 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. Grade x year FEs are indicators for each unique year-grade combination. School x year FEs are indicators for each unique school-year combination. All columns also include a family FE. At the bottom of column (1) we report the coefficient and standard error for the variable *Foreign-born exposure* in the first stage of the 2SLS estimation. Robust standard errors in parenthesis clustered by cohort. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

	Math standardized score (3rd-10th grade)				
	Sample restriction: Lunch Status = No Free/Reduced-price				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	-0.424***	-0.392***	-0.271***	-0.020	-0.065
	(0.067)	(0.065)	(0.062)	(0.082)	(0.117)
Immigrant performance index (Behavior)	-0.107***	-0.105***	-0.096***	-0.080***	-0.088***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
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
Observations	587,588	587,588	586,877	587,588	587,588
R-squared	0.219	0.236	0.271	0.675	0.769
Mean LHS	0.480	0.480	0.480	0.480	0.480
SD LHS	0.867	0.867	0.867	0.867	0.867
Mean RHS	0.0624	0.0624	0.0624	0.0624	0.0624
SD RHS	0.0518	0.0518	0.0518	0.0518	0.0518
Standardized Coefficient	-0.0253	-0.0234	-0.0162	-0.00122	-0.00386

Table A13_A: This table shows the estimates of a linear regression of test scores in mathematics 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, a cumulative index of foreign-born peers' behavioral performance (computed as the average performance index across previous school-specific cohorts including the current grade), and several controls. A school-cohort index of foreign-born behavioral performance is 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 regressions are run on an unbalanced longitudinal sample of US-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. *** p<0.01, ** p<0.05, * p<0.1.

	Math standardized score (3rd-10th grade)				
	Sample restriction: Lunch Status = Free/Reduced-price				
	(1)	(2)	(3)	(4)	(5)
Foreign-born Exposure	0.376***	0.291***	0.305***	0.442***	0.377***
	(0.053)	(0.050)	(0.049)	(0.076)	(0.107)
Immigrant performance index (Behavior)	-0.106***	-0.099***	-0.091***	-0.050***	-0.064***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
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
Observations	691,413	691,413	689,662	691,413	691,413
R-squared	0.251	0.280	0.293	0.627	0.739
Mean LHS	-0.301	-0.301	-0.301	-0.301	-0.301
SD LHS	0.952	0.952	0.952	0.952	0.952
Mean RHS	0.0641	0.0641	0.0642	0.0641	0.0641
SD RHS	0.0522	0.0522	0.0522	0.0522	0.0522
Standardized Coefficient	0.0206	0.0160	0.0167	0.0243	0.0207

Table A13_B: This table shows the estimates of a linear regression of test scores in mathematics 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, a cumulative index of foreign-born peers' behavioral performance (computed as the average performance index across previous school-specific cohorts including the current grade), and several controls. A school-cohort index of foreign-born behavioral performance is 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 regressions are run on an unbalanced longitudinal sample of US-born students eligible for free or reduced-price lunch 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. 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. *** p<0.01, ** p<0.05, * p<0.1.

	Math standardized score (3rd-10th grade)				
	(1)	(2)	(3)	(4)	(5)
		Segregation: current school		Segregation: first school	
		High	*Low*	*High*	*Low*
Foreign-born Exposure		0.296	0.368***	0.213	0.304**
		(0.198)	(0.134)	(0.201)	(0.151)
Foreign-born Exposure (weighted)	0.264***				
	(0.096)				
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
Observations	1,347,103	681,801	656,548	674,090	670,528
R-squared	0.769	0.849	0.858	0.859	0.859
Mean LHS	0.050	0.032	0.078	0.089	0.012
SD LHS	0.993	0.984	0.997	0.975	1.010
Mean RHS	0.037	0.042	0.079	0.052	0.069
SD RHS	0.038	0.036	0.060	0.042	0.060
Standardized coefficient	0.010	0.011	0.022	0.009	0.018

Table A14: This table shows the estimates of a linear regression of test scores in mathematics 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. In Column 1, share of foreign-born students is weighted by a segregation index computed at the school-specific cohort level. In Columns 2 to 5 we estimate the same model as in Table 4 (Column 5), except that we divide the sample based on the segregation index being above or below the median. In Columns 2 and 3 the segregation index is computed in the current school, while in Columns 4 and 5 the relevant segregation index is the one of the initial school in which the student is observed in our data. See the text for details about the construction of the segregation index. All regressions are run on an unbalanced longitudinal sample of US-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. Robust standard errors in parenthesis clustered by school-cohort. *** p<0.01, ** p<0.05, * p<0.1.

	Math standardized score (3rd-10th grade)		
	(1)	(2)	(3)
Diversity dimension:	**Race**	**Race**	**Country**
Population diversity:	All	Foreign-Born	Foreign-Born
Foreign-born Exposure	0.225***	0.228***	0.210***
	(0.076)	(0.077)	(0.079)
Diversity Exposure	0.009	0.008	0.013
	(0.030)	(0.011)	(0.010)
Individual Controls	X	X	X
School x Year FE	X	X	X
Grade x Year FE	X	X	X
Family x Year FE	X	X	X
Observations	1,347,289	1,318,366	1,318,366
R-squared	0.769	0.772	0.772
Mean LHS	0.0504	0.0516	0.0516
SD LHS	0.993	0.994	0.994
Mean RHS	0.0604	0.0618	0.0618
SD RHS	0.0523	0.0521	0.0521
Standardized Coefficient	0.0118	0.0120	0.0110

Table A15: This table shows estimates from three specifications analogous to Column (5) in Table 4, including as explanatory variables three different proxies of diversity exposure calculated as the average exposure across previous school-specific cohorts including the current grade. In Column 1, the diversity index is 1 minus the Herfindahl-Hirschman Index (HHI) for the overall sample of students computed on different reported racial groups. In Column 2 and 3, the diversity index is 1 minus the Herfindahl-Hirschman Index (HHI) calculated only on the sample of foreign-born students based on race (Column 2) and country of origin (Column 3). Robust standard errors in parenthesis clustered by school-cohort. *** p<0.01, ** p<0.05, * p<0.1.

Figures

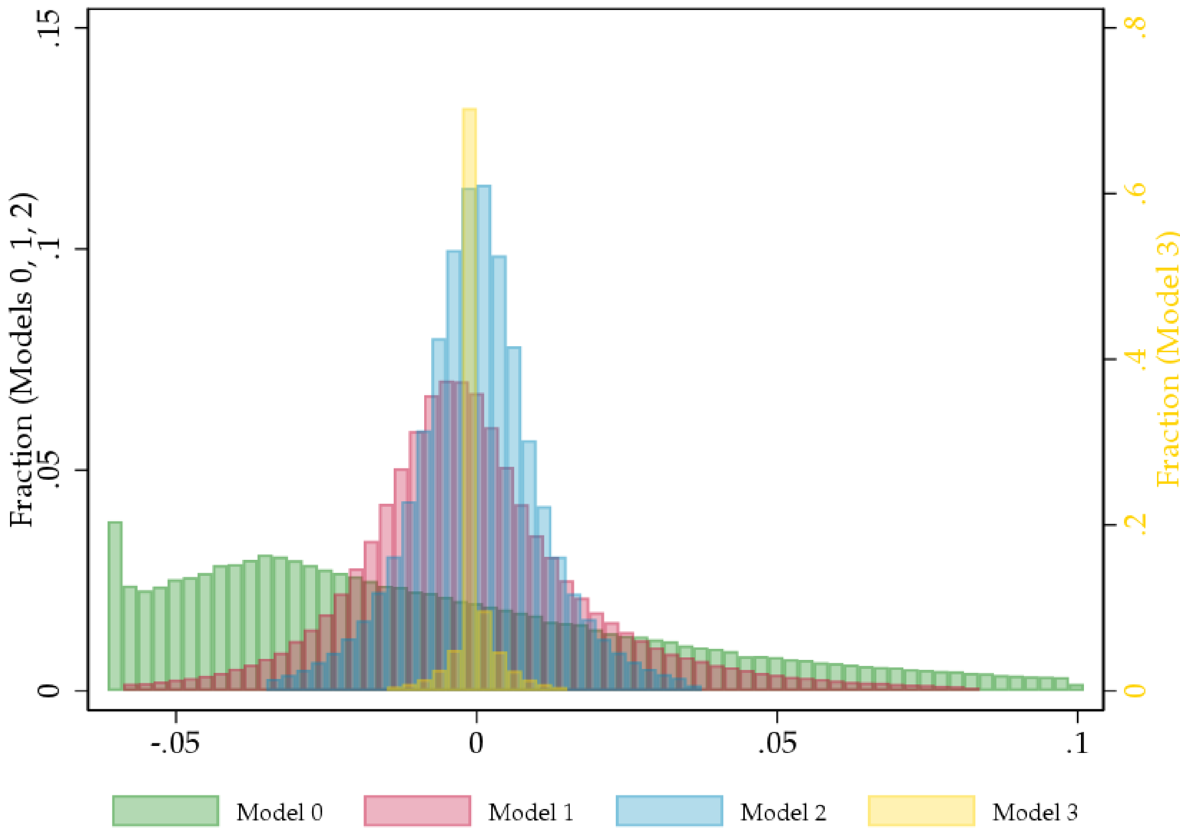

Figure A1: Distribution of cumulative exposure to foreign-born students and its residuals. Model 0 refers to the demeaned distribution (i.e., the raw distribution centered at zero). Model 1 is the distribution of residuals after conditioning on school-year and grade-year Fixed Effects; Model 2 is the distribution of residuals after conditioning on school-year, grade-year, and family Fixed Effects; Model 3 is the distribution of residuals after conditioning on school-year, grade-year, and family-year Fixed Effects. Distributions corresponding to models 0 through 2 are described by the left y-axis, while the distribution corresponding to Model 3 is described by the y-axis on the right-hand side of the graph.

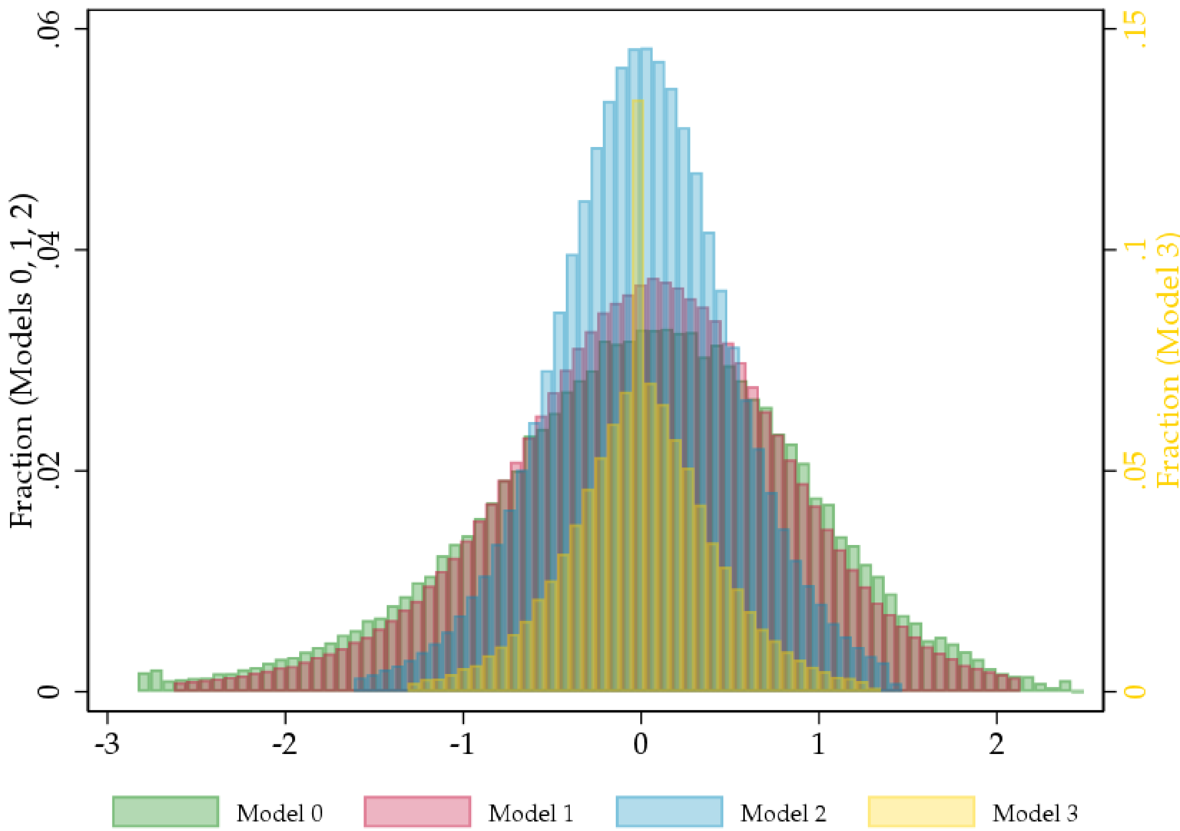

Figure A2: Distribution of standardized math scores and its residuals. Model 0 refers to the demeaned distribution (i.e., the raw distribution centered at zero). Model 1 is the distribution of residuals after conditioning on school-year and grade-year Fixed Effects; Model 2 is the distribution of residuals after conditioning on school-year, grade-year, and family Fixed Effects; Model 3 is the distribution of residuals after conditioning on school-year, grade-year, and family-year Fixed Effects. Distributions corresponding to models 0 through 2 are described by the left y-axis, while the distribution corresponding to Model 3 is described by the y-axis on the right-hand side of the graph.

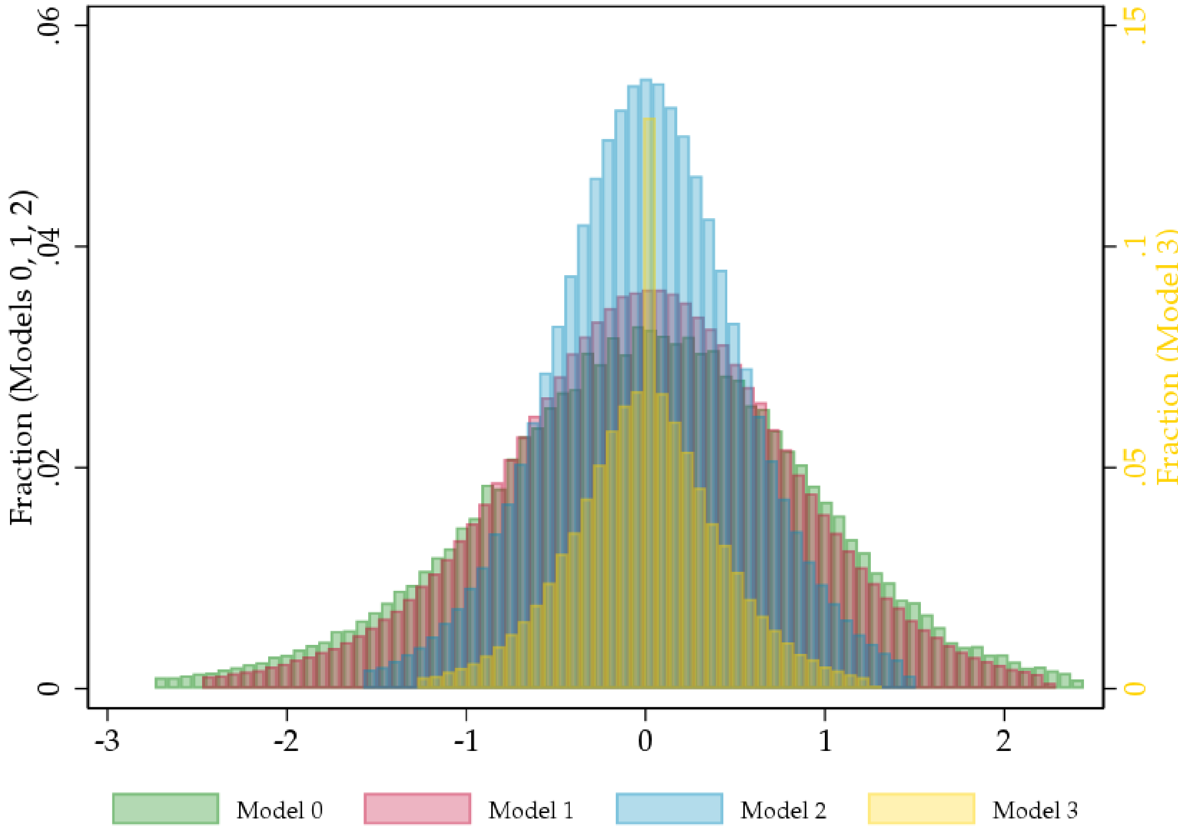

Figure A3: Distribution of standardized reading scores and its residuals. Model 0 refers to the demeaned distribution (i.e., the raw distribution centered at zero). Model 1 is the distribution of residuals after conditioning on school-year and grade-year Fixed Effects; Model 2 is the distribution of residuals after conditioning on school-year, grade-year, and family Fixed Effects; Model 3 is the distribution of residuals after conditioning on school-year, grade-year, and family-year Fixed Effects. Distributions corresponding to models 0 through 2 are described by the left y-axis, while the distribution corresponding to Model 3 is described by the y-axis on the right-hand side of the graph.

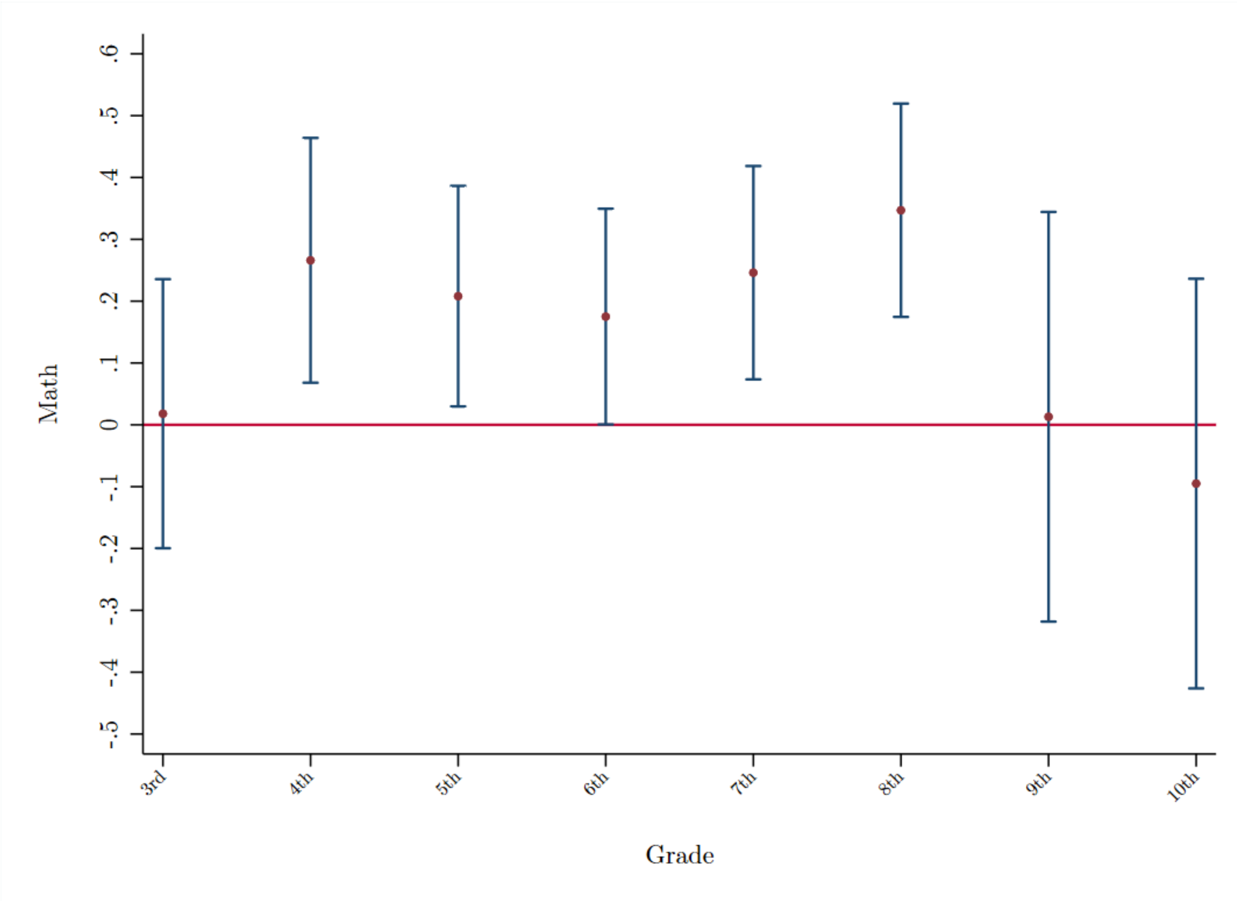

Figure A4: This figure plots the coefficient of the variable *Foreign-born Exposure* in a regression with the same specification as Table 4, column (5), but on the subsample of students enrolled in a given grade.

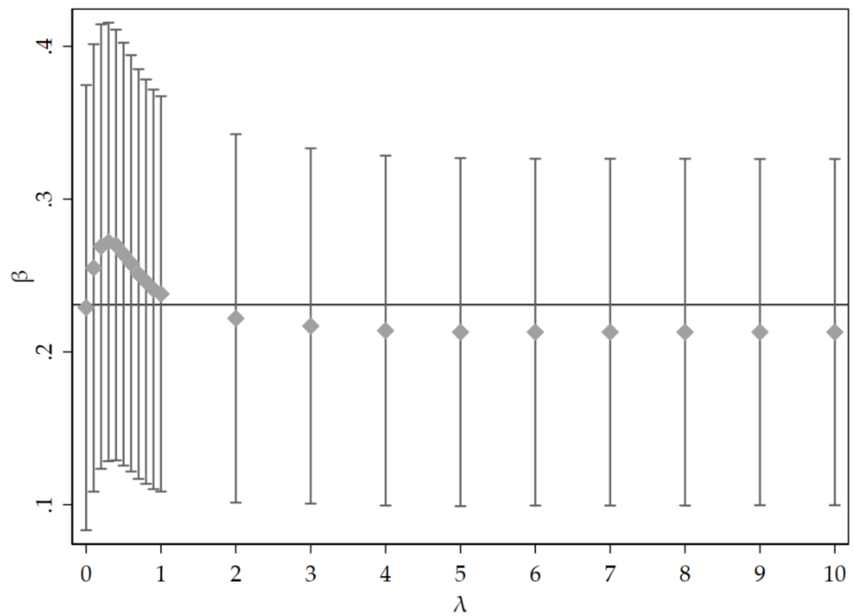

Figure A5: This figure plots the coefficient for cumulative exposure for different lambda based on the equation (1) in the text.

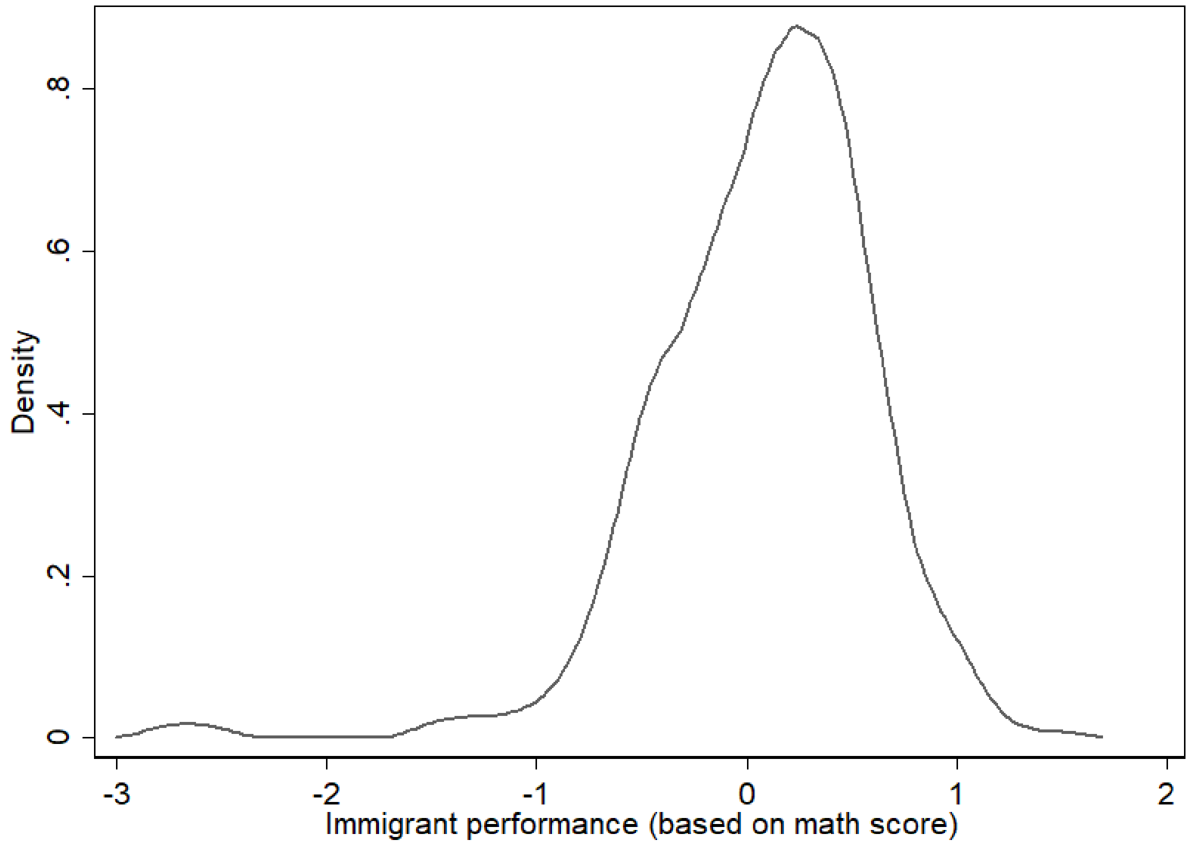

Figure A6A: This graph shows the distribution of the cross-country immigrant performance index based on math performance. We construct the country-specific performance index by averaging the standardized math score by country of origin.

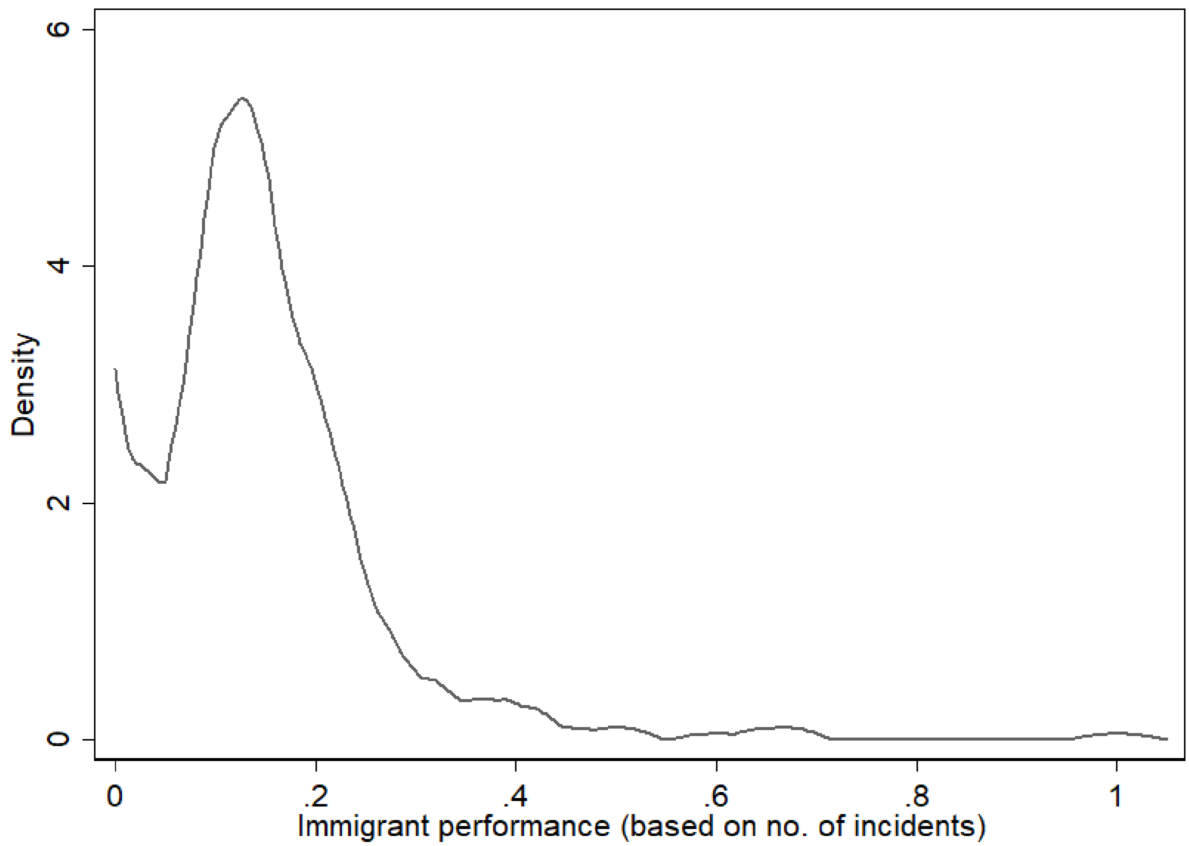

Figure A6B: This graph shows the distribution of the cross-country immigrant performance index based on number of incidents. We construct the country-specific performance index by averaging the number of incidents by country of origin.

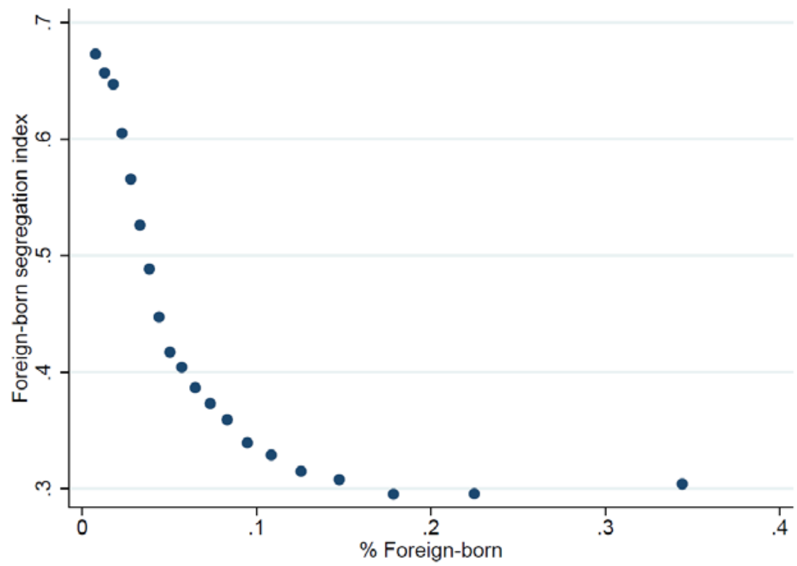

Figure A7: This figure reports the binned scatter plot depicting the average segregation index as a function of the share of foreign-born students across school-grade-year cells. See the text for details about the construction of the segregation index.

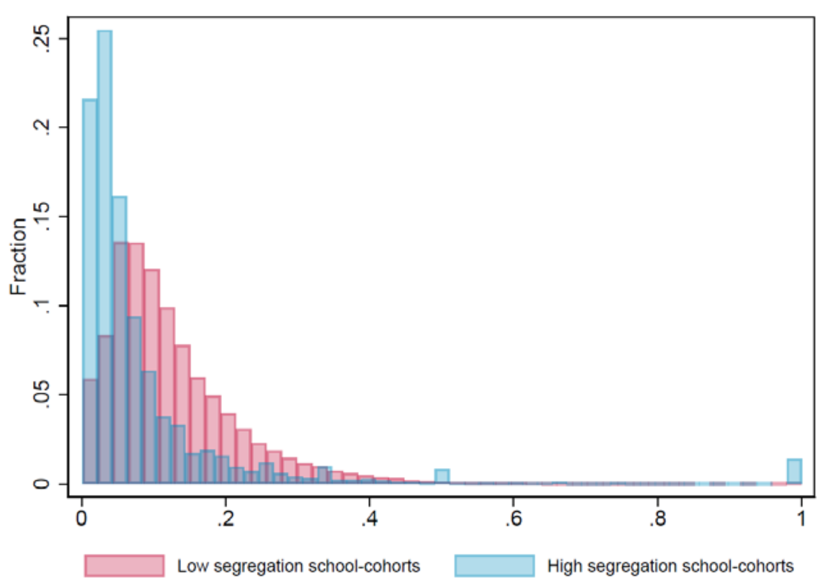

Figure A8: This figure reports the distribution of the share of foreign-born students across school-grade-year cells for cells above and below median segregation index. See the text for details about the construction of the segregation index.