I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11682

Asian Segregation and Scholastic Achievement: Evidence from Primary Schools in New York City

Rocco d'Este
Elias Einiö

JULY 2018

DISCUSSION PAPER SERIES

IZA DP No. 11682

Asian Segregation and Scholastic Achievement: Evidence from Primary Schools in New York City

Rocco d'Este
University of Sussex and IZA

Elias Einiö
VATT and CEP, London School of Economics

JULY 2018

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Asian Segregation and Scholastic Achievement: Evidence from Primary Schools in New York City*

This paper examines the effects of Asian segregation on students' academic performance in New York City primary schools. We use exogenous variation in the share of Asian students across cohorts and schools stemming from a fertility shock among Asian population in the Chinese year of the Dragon. A one-percentage-point increase in Asian student share reduces non-Asian math and ELA scores by 0.03 and 0.05 standard deviations. The effects are largest among black and Hispanic students. We find little evidence of effects among white students. The findings suggest that desegregation policies may generate net benefits in terms of student achievement.

JEL Classification: I20, I29, J15

Keywords: Asian students, student composition, primary education, public schools, test score, student achievement, Chinese Dragon year, race, ethnicity, segregation

Corresponding author:
Rocco d'Este
Department of Economics
University of Sussex
Brighton, BN1 9RH
United Kingdom
E-mail: R.D-Este@sussex.ac.uk

* We are grateful to Jordi Blanes i Vidal, Maria De Paola, Michael Gilraine, Hessel Oosterbeek, Francis Kramarz, Victor Lavy, Devin Pope, Fabian Waldinger, Roberto Weber, and several seminar participants for helpful comments and suggestions.

1 Introduction

Since the famous 1954 Supreme Court case *Brown vs. Board of Education*, which deemed segregation in public schools unconstitutional, extensive school integration efforts have been undertaken in the U.S. public-school sector. Many of these policies have focused on reducing segregation of black students. However, in recent decades, the number of nonblack minority students has grown rapidly in many U.S. school districts, which has led some scholars to argue that the increase in their share could have affected students' academic outcomes in these areas (e.g., Rivkin and Welch, 2006). Despite the growing importance of nonblack minority students, surprisingly little evidence of the impacts of their segregation on scholastic outcomes exists.

In this study, we look beyond white-black school segregation by examining the extent to which Asian students affect their non-Asian peers' academic performance. Asian students are a specifically interesting minority group for at least two reasons. First, they perform relatively well, on average, on academic tests, compared to many other minority groups. Second, child education attitudes in many Asian cultures, especially in the Chinese culture and other Southeast Asian cultures that it influences, put specific emphasis on competitiveness and disciplinary sanctions to incentivize children's schoolwork (e.g., Hesketh et al., 2010). Arguably, the increasing share of Asian students might have different impacts on scholastic outcomes from changes in the shares of other minority groups.

The key econometric challenge in identifying the causal effects of the segregation of Asian students on their peers' scholastic outcomes arises from the potential selection of Asian children into schools. Public-school systems typically allocate students to schools by place of residence, which might be determined by several characteristics that could also affect children's scholastic outcomes but are not observed by researchers. For instance, parents who are more concerned with their children's education could place them in better schools, often located in more expensive residential neighborhoods (e.g., Black, 1999; Black and Machin, 2011). In such neighborhoods, peer quality could also be higher on average. Therefore, if Asian parents tend to choose better neighborhoods, the fraction of Asian students will likely be positively correlated with unobserved school and peer quality, which would induce a positive bias in regressions of peer performance on the share of Asian students.

We address these issues by exploiting plausibly exogenous variation in the share of Asian students stemming from the common belief among the Asian population that children born in the Dragon year are luckier and brighter than those born under other zodiac signs. Prior evidence indicates that this belief generates considerable positive shocks to fertility in Asian populations in the Dragon years.[4] Our research setting exploits the fact that the effect of these fertility shocks on racial composition varies geographically with the Asian population's local relative size. In areas with a small historical share of Asian population, the Asian fertility shock in the Dragon year induces only small differences in the share of Asian students between cohorts, but in areas with a large Asian population share, it induces a disproportionately large share of Asian students in the Dragon cohort compared to other cohorts. These differential effects allow us to construct instruments based on the historical Asian population share to study the impact of the share of Asian students on their non-Asian peers' test scores.

Our study covers 1,081 public primary schools in New York City (NYC), which has one of the largest Asian populations among U.S. metropolitan areas (around 14% of the population). We use NYC Department of Education data on average public-primary-school math and English language arts (ELA) test scores in third through eighth grade by school, year, and race/ethnicity. We geocode schools by address and link them to 1990 census tract-level population data to measure the historical racial/ethnic population structure in a school's neighborhood. In our analysis, the key Asian groups are those that are influenced by the Chinese culture. For this reason, we base our preferred instrument on the historical local Chinese population share, but we also provide results for instruments based on wider Asian groups. Throughout our analysis, we include fixed effects by school, cohort, grade-by-year, and race/ethnicity.

We start by examining the fertility shock's impact on the number of Asian students in the Dragon cohort and other cohorts across areas with high and low historical Chinese population shares. We show that a 10-percentage-point increase in the historical Chinese population share corresponds to an approximate 1.2 additional Asian students at the grade level in the Dragon cohort compared to the number of Asian students in other cohorts within a school (around an 8.6% increase from the average number of Asian students in third grade). We detect no corresponding statistically significant effects of the instrument

[4] Mocan and Yu (2017) showed that births spike in the Dragon years 2000 and 2012 in China. Johnson and Nye (2011) provided evidence of the Dragon effect for the 1976 cohort among Asian immigrants in the United States. Yip et al. (2002) documented the dragon effect in Hong Kong for cohorts born in 1976, 1988, and 2000.

on the number of non-Asian students. These findings indicate that schools located in areas with large 1990 Chinese populations are not fully constrained in their capacities and can enroll additional Asian students without significantly reducing the number of students in other racial/ethnic groups.

Our instrumental variable (IV) analysis detects a statistically significant negative impact of the share of Asian students on the non-Asian math and ELA test scores ($p<0.05$ for both). Our preferred estimates suggest that a one–percentage-point increase in the share of Asian students reduces the non-Asian math score by around 0.03 standard deviations (henceforth σ) and the non-Asian ELA score by around 0.05 σ. The negative effect on ELA scores is mainly driven by adverse effects on Hispanic and black students and schools where the non-Asian student population is ethnically/racially more fractionalized. We identify some differential impacts on students' proficiency distributions across subjects. Most notably, we detect a statistically significant and sizeable positive effect on the share of poorly performing non-Asian students who are well below proficiency in math, which suggests that a larger share of Asian students might cause a larger share of their peers to lag behind in math.

We find no evidence of changes in the number of teachers per student, teacher quality, class size, congestion, or student attrition generating the results. Our estimates are little affected by these factors likely because institutional constraints limit variation in class size; school funding is primarily based on enrollment, which we control for in our IV analysis; and school transfers are granted only in rare circumstances. Overall, our findings are consistent with the interpretation that the negative effects on non-Asian test scores are caused by the larger share of Asian students.

The impacts of the increasing share of Asian students on non-Asian achievement might operate through several channels. First, the adverse effects on non-Asian test scores could be related to peer effects between students. For instance, additional Asian students, who are on average well-performing, might discourage non-Asian students, lowering their study motivation and leading them to exert less effort. Second, a change in student composition could affect scholastic outcomes through teacher responses (e.g., Duflo et al., 2011; Lavy et al., 2012). Additional Asian students might reduce the teacher's attention available to non-Asian students and affect the teacher's curriculum in terms of coverage and pace. For instance, our finding that Asian students, who are on average extremely well-performing in math, increase the share of poorly performing non-Asian students in this subject might be generated by an increase in the pace and coverage of math instruction. The finding of asymmetric impacts on ELA and math proficiency distributions suggests that some of the mechanisms through which racial/ethnic segregation affects scholastic performance are likely to be subject-specific.

This study contributes to literature examining the impacts of racial composition on scholastic achievement. Much of the previous research has examined the impacts of black segregation (e.g., Cook and Evans, 2000; Angrist and Lang, 2004; Hanushek and Rivkin, 2009; Hanushek et al., 2009). A notable exception is Hoxby (2000). In her study examining the impacts of student composition on test scores in public primary schools in Texas, she employs idiosyncratic, cohort-to-cohort changes in student composition of a grade level within a school. She finds some evidence of positive impacts of the share of Asian students on non-Asian math scores and little evidence of impacts on reading scores. Contrary to her findings, we find significant negative effects of the share of Asian students on non-Asian math and ELA scores. Our varied findings could stem from the fact that the share of the Asian population and spatial variation in it are substantially larger in NYC than in Texas. Moreover, our research setting induces substantial quasi-experimental variation in the share of Asian students, stemming from the fact that the 1990 local share of the Chinese population varies widely across school neighborhoods in NYC and is strongly correlated with the fertility shock in the Dragon year 10 years later. Hence, our IV analysis is likely to recover the impact of a larger shift in Asian segregation than previous studies.

More generally, our study is linked to the literature on peer effects in education.[5] In particular, our identification strategy is related to previous studies employing quasi-experimental variation in peer composition arising from explicit shocks to the local population structure (e.g., Imberman et al., 2012).[6] Because the shock on the share of Asian children in the Year of the Dragon is not restricted to the school environment, our identification strategy could prove useful in estimating the impacts of Asian segregation on a variety of other economically and socially relevant outcomes. The relevance of assessing such impacts is growing in importance in the wake of China becoming the major immigrant-sending region in the U.S. (Jensen et al., 2015).

The paper proceeds as follows. Section 2 provides details of the institutional background and presents data sources and descriptive statistics. Section 3 documents the

[5] See Sacerdote (2011) for a comprehensive survey.

[6] See also Ballatore, Fort, and Ichino (forthcoming), who exploited class size rules to identify the impacts of increasing shares of immigrant students on native test scores in Italy, and Geay, McNally, and Telhaj (2013), who examined the impacts of the share of nonnative English speakers on scholastic achievement in England.

shock on the number and share of Asian students in NYC schools induced by our instrument and presents the estimation strategy. Section 4 provides the main results and numerous robustness checks verifying the empirical design's validity. We also investigate heterogeneity of the effects by grade, race/ethnicity, and school characteristics. Section 5 examines the extent to which school-level responses and other potential mechanisms, such as attrition of students, could explain our findings. Section 6 concludes.

2 Institutional Background and Data

2.1 Data on New York City Public Primary Schools

The NYC Department of Education (DOE) is one of the largest schooling authorities in the U.S., serving around 1.1 million students. In principle, the DOE provides a place in a local public school for all children the year they turn five (kindergarten). The allocation of students to public primary schools is based on *parents' residential addresses,* and only special needs and circumstances could enable students to move to an undesignated school.[7] These special circumstances are: 1) medical reasons, 2) the student's safety, 3) parent's employer being located far from the designated school, 4) a sibling attending a different school, and 5) own school being listed as a school in need of improvement or low-achieving school in the last two years.[8]

Test scores, proficiency groups, and grade-level average class size. Our analysis uses data on 1,081 public primary schools that reported the results of the New York State English language arts (ELA) and math tests for third through eighth grade in the academic years 2005/2006 through 2011/2012. We use publicly available test score files for math and ELA mean test scores by school, grade, year, and race/ethnicity (white, black, Asian, and Hispanic) provided by the DOE.[9] To improve comparability across grades and years, the DOE adjusts raw test scores with a scaling procedure that accounts for question difficulty, their capacity to differentiate between high- and low-performing students, and the likelihood of getting a correct answer by guessing. Throughout our analysis, we further account for unobserved test heterogeneity across grades and years by controlling for grade-by-year fixed effects. The test score files also include information on the number of

[7] Each school district's Community Education Council sets the boundaries for school zones.
[8] Title I of the No Child Left Behind Act.
[9] The data can be accessed online at:
http://schools.nyc.gov/NR/exeres/05289E74-2D81-4CC0-81F6-E1143E28F4C4,frameless.htm
The tests are administered in the spring semester. The 2006 test score, for instance, is for the academic year 2005/2006. To avoid disclosing individual scores, the DOE provides the ELA and math mean scores only when the number of students in a school-grade-race/ethnicity-year cell is larger than four.

students attending the test sessions and the number of students at each of four proficiency levels. The score thresholds for proficiency levels are determined annually for each grade at the state level by a panel of experts. The four proficiency levels are well below proficient, below proficient, proficient, and above proficient. The DOE provides also information on average class size at the school-grade-year level, but this data is unavailable for the academic year 2005/2006. Hence, it can be included in estimations using a slightly restricted sample. We provide further details of these data in appendix A.1.

Enrollment. We use information on the number of students taking the ELA and math tests to measure enrollment by school, grade, year, and race/ethnicity. ELA and math tests are obligatory for all students; therefore, the number of students attending these tests can be expected to be fairly close to actual enrollment. To reduce variation due to the possibility that some students cannot attend all tests, we use the larger value if attendance in the two tests within a school-grade-year-ethnicity/race cell is not equivalent. We use this measure of enrollment to construct student shares by ethnicity/race at the school-grade-year level. We examined the enrollment measure's accuracy by aggregating it at the school-year level and comparing it to the corresponding annual school-level enrollment figures drawn from New York State School Report Cards. The correlation between these variables is 0.98, which indicates that test attendance provides a reasonably good measure of enrollment.

School-level variables. We link the test score files to annual school-level variables drawn from the School Report Cards, which provide information on the number of teachers, total number of classes taught, suspensions, the fraction of students with reduced-price or free lunch status, and class size (school-level average across first through sixth grade).[10] We use the latter variable, covering all years in our test score data, as a complementary measure of class size. The data also include information on the number of teachers with valid teaching certificates and with less than three years of experience, which we use to construct measures of teacher quality. We provide further details of the report-card data in appendix A.2.

Geocoding of schools. The report cards provide school addresses, which were used to geocode schools. Our primary source for school coordinates is the U.S. Census Batch Geocoder.[11] The resulting address matches and coordinates were manually checked, and schools with missing coordinates were manually geocoded. We assigned coordinates to

[10] The data are available at https://data.nysed.gov/downloads.php.

[11] https://geocoding.geo.census.gov/geocoder/geographies/addressbatch?form

around 99% of schools.[12] Appendix A.3 provides further details of the geocoding procedure.

Local population structure. We constructed variables for the historical ethnic/racial population structure in a school's neighborhood from 1990 tract-level census data on population by ethnicity/race and census-tract boundary shapefiles provided by the Minnesota Population Center.[13] We used a GIS procedure to find census tracts within 500 meters of the school. When several nearby census tracts were identified, we used the population-weighted average of ethnic population shares. We call the area covered by these nearby census tracts the school neighborhood. The 1990 Chinese population share in a school neighborhood is our primary measure of historical Chinese exposure. We also report results using other measures based on a wider set of Asian groups (e.g., Asians excluding Asian Indians; all Asians) and various geographic scopes (census tracts within 1,000, 2,000, and 3,000 meters). Further details of the construction of these variables are provided in appendix A.4.

Descriptive statistics. We restrict data to school-grade-year-race/ethnicity cells for which both math and ELA mean test scores are observed. Thus, we have the same schools and student groups for both outcomes in each specification. Table 1 provides summary statistics for the estimation sample and for two subsamples where the 1990 Chinese population share in the school neighborhood is below and above the median. Panel A displays weighted means and standard deviations of the mean test scores for all students and by race/ethnicity using the number of students attending the test as a weight. The mean math (ELA) score is 675.1 (658.6) with a standard deviation of 23.0 (16.5). Asian students have the highest mean in math, followed by white students, for whom it is around a half standard deviation lower. Asian and white students score, on average, almost similarly in the ELA test. In both subjects, Asians have the largest share of students at the highest proficiency level – around 48% in math and 12% in ELA – and the smallest share of students at the lowest proficiency level – around 2.5% in math and 4.4% in ELA (see appendix table A2). Overall, Asian primary-school students are high-achieving, and only a very small fraction of them score poorly in the math and ELA tests. Black and Hispanic students score significantly worse in both tests, on average, than white and Asian students.

[12] Appendix figure A1 shows an example of geocoded schools and census tracts in Manhattan.
[13] We used the file nhgis0001_ds120_1990_tract.

Black students perform slightly poorer in math than Hispanic students, and these two groups perform similarly in ELA.[14]

Panel B of table 1 shows descriptive statistics for enrollment and shares of students by race/ethnicity at the school-grade-year level. Mean enrollment at the grade level is around 125 students. Hispanic students have the highest average share (39.6%), followed by black (32.2%), white (14.9%), and Asian (13.3%) students. In school neighborhoods that had a Chinese population share above the median in 1990, Asians make up around 24% of students, and in schools below the median, the Asian share is around 4.3%, on average. Grade-level average class size is 24.8 students. Appendix table A3 provides descriptive statistics for school characteristics.

3 Estimation Strategy

3.1 The Fertility Spike in the Dragon Year

In the Chinese calendar, the Dragon year appears once every 12 years. According to a widespread belief among many East Asian cultures, children born in the Dragon years are luckier, brighter, and more likely to flourish. This belief generates fertility shocks in the Dragon years in populations among which this belief is prevalent. Previous research finds considerable spikes in birth rates in the Dragon years in China (Mocan and Yu, 2017), many East Asian regions (e.g., Goodkind, 1995; Yip et al., 2002), and among Asian populations in the U.S. (Johnson and Nye, 2011).

Our empirical strategy employs variation in the share of Asian students due to the Dragon year starting on February, 5 2000, and ending on January 23, 2001. Although the Chinese calendar year does not perfectly overlap with the Western calendar year, most children born in this Chinese Dragon year are also born in the Western year 2000.[15] Asian births per 1,000 individuals were around 7.5% higher in the U.S. in 2000, compared to the average rate in the years 1998-1999 and 2001-2002 (U.S. Census Bureau, 2013).

Because birth rates among other ethnic/racial groups deviated only a little in 2000, the fertility shock among Asian population increases the share of Asian children in the 2000 birth cohort.[16] As a result, the share of Asian primary-school students in the 2000 birth

[14] Appendix figure A2 shows test score distributions by race/ethnicity.
[15] The fact that the Chinese Dragon year overlaps with the year 2001 by 23 days may raise the concern that Dragon children born in the first 23 days of 2001 affect our results. However, we show that excluding them from the sample has little impact on our estimates. Therefore, hereafter, we use the phrases “year 2000” and “Dragon year” interchangeably and refer to the cohort born in the Western year 2000 as the Dragon cohort.
[16] Appendix figure A3 shows annual birth rates per 1,000 individuals by race/ethnicity in the U.S. over the years 1996-2004.

cohort compared to other birth cohorts is also affected. Figure 1 displays average Asian enrollment for birth cohorts 1992 through 2003 in the 1,081 public primary schools observed in our data. It shows a sharp spike in the number of Asian students in the 2000 cohort. The average Asian enrollment increases from 9,610 students in the 1999 cohort to 10,520 students in the 2000 cohort. The number of Asian students is also higher in the 2000 cohort than in the 2001 and 2002 cohorts, and it is around 7.5% higher than the average number of Asian students in cohorts born in 1998-1999 and 2001-2002.[17] In what follows, we explain how we exploit geographic variation in the magnitude of this fertility shock stemming from variation in the historical local share of Asian population to identify the causal impacts of the share of Asian students on their non-Asian peers' test scores.

3.2 Geographic Variation in the Size of the Fertility Shock in the Dragon Year

Our empirical strategy exploits the fact that the effect of the fertility shock in the Dragon year on racial composition varies geographically with the local historical share of Asian population. To see the link between the historical share of Asian population and the fertility shock in the Dragon year, consider a school neighborhood with A Asian and H non-Asian births in a cohort born in a non-Dragon year. The share of Asian children in a non-Dragon cohort is then $a = A/(A + H)$. Suppose that the Dragon year increases Asian births by $\delta \cdot 100\%$ and has no impact on the number of non-Asian births. The local share of Asian children in the Dragon cohort will then be $a_D = (1 + \delta)A/((1 + \delta)A + H)$. It is straightforward to show that the increase in the fraction of Asian children between the Dragon and non-Dragon cohorts, $a_D - a$, is the following function of the size of the fertility shock in the Dragon year and the share of Asian children in the non-Dragon cohorts:

$$g(a, \delta) = \frac{a - a^2}{\delta^{-1} + a}. \tag{1}$$

This function is concave and nonnegative when $a \in [0,1]$ and has a maximum at $a = 0.5$ for $\delta > 0$. This relationship implies that, for instance, between areas with $a = 0$ and $a = 0.5$, a fertility shock of $\delta = 0.075$ in the Dragon year induces a relative increase of around 1.8 percentage points in the share of Asian children in the Dragon cohort compared to other cohorts.

[17] It is worth noting that we will control for cohort fixed effects throughout our regression analysis, which accounts for the positive trend in the number of Asian students observed in figure 1.

Figure 2 shows the function $g(a, 0.075)$ for $a \in [0,1]$ and the empirical histogram of the share of Asian students observed in our data in the third grade in birth cohorts born one to three years before the Dragon cohort (i.e., in 1997-1999). The figure indicates that the share of Asian students in these cohorts is within the range where $g(\cdot, 0.075)$ is increasing for the majority of schools. As a result, the average derivative of $g(\cdot, 0.075)$ evaluated across the distribution of the share of Asian students is positive (0.055). Although the nonlinear theoretical relationship in equation (1) could affect our analysis, our results are actually not very different when we account for it in the IV estimations. Therefore, we use the simpler linear-instrument specification throughout our analysis, but we also report results for overidentified specifications allowing for a nonlinear first stage.

3.3 IV Estimation

In our analysis, the key Asian groups are those that are influenced by the Chinese culture. For this reason, we base our preferred instrumental variables on the local Chinese population share in 1990. We exploit variation in the share of Asian students induced by the disproportionately large fertility shock in the Dragon year in areas with high historical Chinese population share by estimating the following IV model:

$$y_{rsgt} = \rho_1 CS_{s,1990} + \rho_2 Dragon_{gt} + \gamma AS_{sgt} + \beta_1' X_{rsgt} + u_{rsgt} \tag{2a}$$

$$AS_{sgt} = \tau_1 CS_{s,1990} + \tau_2 Dragon_{gt} + \tau_3 CS_{s,1990} \times Dragon_{gt} + \beta_2' X_{rsgt} + \nu_{sgt} \tag{2b}$$

where y_{rsgt} is the mean test score of non-Asian students in racial/ethnic group r in school s, grade g, and year t. AS_{sgt} is the share of Asian students in school s, grade g, and year t. $Dragon_{gt}$ is a binary indicator taking the value 1 if the cohort born in 2000 is in grade g in year t and 0 otherwise. $CS_{s,1990}$ is the Chinese population share in 1990 in the neighborhood of school s. X_{rsgt} is a vector of control variables. In our baseline specification, we include school, cohort, race, and grade-by-year fixed effects.[18] We weight regressions by the number of students taking the test and cluster standard errors at the census-tract level.

The inclusion of school fixed effects controls for time-invariant differences across schools. A school fixed-effects regression of equation (2a) would identify the parameter γ from within-school variation in the share of Asian students. However, unobserved shocks

[18] Note that when school and cohort fixed effects are included, the main effect terms for $CS_{s,1990}$ and $Dragon_{gt}$ become redundant.

to school-level variables that affect the share of Asian students and non-Asian test scores could bias the school fixed effects estimate. For instance, if Asian families tend to move into neighborhoods of schools where test scores are expected to improve more than in other schools, such a selection might bias the estimates of γ obtained from school fixed effects regressions.

To account for such biases, the IV model uses the interaction term between the Dragon cohort dummy and the Chinese population share in 1990 in a school's neighborhood as an instrument for the share of Asian students in the school over 15 years later (2006–2012).[19] The first-stage coefficient on the instrument, τ_3, recovers the difference in the effect of the fertility shock on the share of Asian students in the Dragon cohort compared to other cohorts between school neighborhoods with high and low historical Chinese population shares. When conditioning on school fixed effects, the identifying variation comes from the disproportionately large share of Asian children in the Dragon cohort, compared to other cohorts within the same school located in a neighborhood with a large historical Chinese population share.[20]

The key identifying assumption in our empirical strategy is that the population shock in the Dragon cohort in areas with a historically larger Chinese population share is uncorrelated with other unobservable factors that affect the share of Asian students and non-Asian test scores. Therefore, the key threat to identification is that non-Asian students in the Dragon cohort perform disproportionately better or worse than students in other cohorts within the same school in areas that had a high Chinese population share in 1990 compared to areas that had a low Chinese population share in 1990. To corroborate our strategy's validity, we show that our results are little affected when we control for year and year of birth trends interacted with the Chinese population share from 1990. This

[19] Alternatively, one could construct instruments based on the share of Asian students in the older non-Dragon cohorts observed in our school data. We prefer instruments based on measures of the historical ethnic/racial population structure retrieved from the census data for two reasons. First, census data allow us to use the share of population at a more detailed level of race/ethnicity, compared to our school data, in which we do not observe subgroups of Asian students. For instance, the census data allow us to construct historical local population share instruments that exclude Asian Indians among whom the beliefs about the Dragon year are not as prevalent. Second, an instrumental variable based on the historical population share provides a more plausible source of exogenous variation in the share of Asian students because it is realized 10 years before the relevant Dragon cohort is born.

[20] The empirical specification is similar to that of Hoxby (2000) and Hanushek et al. (2009) because it exploits variation in student composition across cohorts and within schools. However, instead of using idiosyncratic variation within schools, we employ variation in student composition stemming from an explicit historical population shock in the school's neighborhood. In this respect, the empirical strategy is related to that of Imberman et al. (2012), who employed shocks to the share of students who are evacuees stemming from the population flows from areas hit by Hurricane Katrina to nearby areas that were not directly affected by the natural disaster.

indicates that neighborhood-specific trends correlated with non-Asian test scores are unlikely to be a major source of bias in our IV analysis.

Figure 3 displays geographic variation in the Chinese population share in 1990 by census tract, on which our preferred instrument is based. NYC is an especially suitable metropolitan area for implementing our research design because the Chinese population share in 1990 varied considerably across neighborhoods. The figure also shows schools in our data. Importantly, we observe several schools in areas with high and low historical Chinese exposure.

3.4 The Effect of the Instrument on Student Enrollment and Composition

We first examine the impact of the disproportionately large Dragon cohort on the enrollment of Asian and non-Asian students. Table 2 reports reduced-form effects of the instrument on Asian and non-Asian student enrollment in third grade, the earliest grade that we observe in our data.[21] The regression controls for school, year, and cohort fixed effects. The instrument has a positive and highly significant effect on Asian third-grade enrollment. A 10-percentage-point increase in the Chinese population share in 1990 in a school's neighborhood induces around 1.2 additional Asian students in the Dragon cohort compared to other cohorts, which corresponds to an increase of around 8.6% from the sample mean of 13.9 Asian students. We detect no statistically significant effects on total enrollment of non-Asian students. The coefficient (standard error) for all non-Asian students is only –0.0136 (0.0363). Also, the coefficients for enrollment of non-Asian subgroups (white, black, and Hispanic) are all insignificant and of small magnitude.[22]

Looking at the results for enrollment of all students in column 6, the coefficient is positive and significant, which suggests that a 10-percentage-point increase in the Chinese population share in 1990 induces a relative increase of around one student in the Dragon cohort. These findings indicate that the increase in Asian enrollment induced by the instrument does not significantly crowd out non-Asian students and, as a result, increases total enrollment in the Dragon cohort.

[21] In section 5 below, we also examine the instrument's impacts on attrition of students in subsequent grades but find little evidence of it.

[22] For instance, the magnitude of the coefficient for Hispanic students, which is the largest among non-Asian subgroup, corresponds to a decrease of around 0.15% up from the sample mean of 39.39 Hispanic students when the Chinese population share in 1990 in a school's neighborhood increases by 10 percentage points. We also show that our results are little affected when we control for the shares of students in non-Asian subgroups.

While the above results suggest that the instrument has a strong impact on Asian segregation, the increase in Asian enrollment could affect scholastic performance of non-Asian students through channels other than racial/ethnic composition. It might be that Asian students in the Dragon cohort born in areas with a large Chinese population share in 1990 might perform differently from other Asian Dragon students residing in less segregated areas. If that were the case, our IV strategy would recover the joint effect of a change in Asian segregation and the additional effect of a change in Asian achievement distribution.[23] To examine this possibility, columns 7 and 8 show the instrument's impact on Asian third-grade test scores. The estimates are small and insignificant for math and ELA.[24] These results indicate that we cannot reject the null hypotheses that the additional Asian students in the Dragon cohort achieve similarly well as students in other Asian cohorts in areas with high and low historical Chinese population shares.

Another potential concern is that the increase in enrollment affects class size—which researchers have shown to negatively affect test scores (e.g., Angrist and Lavy, 2009; Chetty et al., 2011; Fredriksson et al., 2013). To examine this possibility, column 9 shows reduced-form effects of the instrument on average class size at the school-grade-year level. The coefficient is small, insignificant, and negative, suggesting that increases in class size are unlikely to drive our IV results.[25] Below we also run specifications controlling for a rich set of school- and grade-level characteristics, including grade-level enrollment and average class size. Reassuringly, adding these controls turns out to have little impact on our results. Overall, the findings suggest that our IV strategy is likely to recover the impacts of an increase in the share of Asian students rather than the impacts of a change in average Asian peer achievement, congestion, or class size.

4 Result

In this section, we present estimates of the impacts that the share of Asian students has on non-Asian students' test scores. We start by reporting the baseline results for all non-Asian students and show that the results are robust against several potential sources of

[23] It is worth noting that because Asian students are, on average, high-achieving, an additional Asian student has, on average, a positive impact on the average peer achievement. We consider this a potential channel through which Asian segregation could affect non-Asian scholastic achievement.

[24] Appendix figure A4 plots Asian test score distributions, comparing the Dragon cohort with cohorts born three years before and after the year 2000. The distributions appear to be very similar for both subjects. Moreover, the impact of the instrument on Asian test scores is small and insignificant in the sample including all grades (appendix table A5).

[25] The corresponding coefficient is also small and insignificant in the sample including all grades (appendix table A5).

confounding variation. We then examine the heterogeneity of the impacts by grade level, ethnic/racial subgroup, and school characteristics.

4.1 Effects on Test Scores

Table 3 shows our baseline results for the IV and the ordinary least squares (OLS) estimates of the effect that Asian-student share has on average non-Asian ELA (panel A) and math (panel B) scores, based on the IV model in equation (2). The specifications in this table control for the school, cohort, ethnicity/race, and grade-by-year fixed effects. The first-stage coefficients are around 0.15 and significant at the 5% risk level. The point estimates indicate that a 10-percentage-point increase in the Chinese population share in a school's neighborhood in 1990 increased the share of Asian students in the Dragon cohort, compared to other cohorts, by around 1.5 percentage points.

The reduced-form estimates are negative and statistically significant for non-Asian ELA and math scores. The IV estimates are –0.765 for ELA and –0.639 for math ($p<0.05$ for both), which suggests that a 10-percentage-point increase in the share of Asian students reduces non-Asian ELA and math scores by around 7.7 and 6.4 points, respectively (around 0.49σ and 0.29σ, respectively). The fourth column of table 3 shows the corresponding OLS estimates. These are positive, small, and insignificant for both subjects. The positive bias in the OLS estimates is in line with the existence of unobserved factors that vary over time within schools and that are positively correlated with both the share of Asian students and non-Asian student achievement.

To provide a benchmark for the magnitude of the estimated effects, we compare them to the results from two other studies in which researchers exploited quasi-experimental variation in primary-school peer composition. Imberman et al. (2012) find in a study of Houston elementary schools that a 10-percentage point increase in the share of students who had escaped Hurricane Katrina reduced the math scores of non-evacuee students in the second and third achievement quartiles by around 0.1σ. Angrist and Lang (2004) find that a 10-percentage-point increase in the share of Metco students reduced the reading and language test scores of black, non-Metco third graders by around 0.6σ. Our estimates fall within the range of these effects.

4.2 Robustness Analysis and Placebo Tests

We next examine the robustness of our results by considering numerous factors that could invalidate the interpretation that the IV coefficient recovers the causal effect that the share of Asian students has on their non-Asian peers' academic performance. The results are shown in table 4.

Nonlinearity of the First Stage. We start by estimating an overidentified specification, including the square of the local Chinese population share in 1990 interacted with the Dragon dummy as an additional instrument (column 2 of table 4). This specification is motivated by the concave function of the theoretical impact in equation (1); it allows for a nonlinear first-stage relationship between the instrument and the share of Asian students. The IV estimates for this specification have a similar magnitude as those in the single-instrument baseline specification, and they are significant at the 5% risk level for both math and ELA.

Trends. One concern is that differential trends in non-Asian test scores between schools with high and low historic Chinese exposure might drive our results. We hence add as a control variable a term for the interaction between the second-order polynomial of the year and the local share of the Chinese population in 1990 (column 3). By allowing for differences in trends based on levels of historical Chinese exposure, the math score estimate increases to –0.949, which is significant at the 5% risk level, and the ELA score estimate reduces to –0.597, which is still significant at the 10% risk level. In column 4, we add similar terms for interactions between the second-order polynomial of birth year. This allows for different trends based on historical Chinese exposure across cohorts. The inclusion of these variables slightly reduces the estimate for math (compared to column 3) but the estimate is larger (compared to the baseline effect) and it is significant at the 10% risk level. For ELA, allowing for cohort trends reduces the precision of the estimation, but the point estimate still has a relatively large magnitude. In column 5, we estimate a specification similar to the one in column 4, but allow for nonlinear first stage. This has little impact on the point estimate for math, although it does slightly improve the estimate's precision. The estimate for ELA is higher than in the corresponding single-instrument specification in column 4 and it is significant at the 10% risk level. Overall, these findings suggest that differences in trends due to levels of historical Chinese exposure do not significantly affect the results.

Congestion. As discussed above, our instrument induces an increase in grade-level enrollment, which raises the concern that congestion caused the negative academic impacts on non-Asian students. Such negative impacts could be due to increased disruption, for instance (e.g., Lazear, 2001; Hanushek et al., 2004). Moreover, having more students could mean that teachers have fewer resources per student, which could adversely affect scholastic outcomes. We examine the extent to which congestion can explain our results by adding control variables for grade size (enrollment at the school-grade-year level). The specification in column 6 of table 4 includes controls for a second-order polynomial of grade size as well as 10 dummies for grade-size decile groups. Controlling for grade size has little impact on the point estimates and slightly improves the precision of the estimation. The point estimates for ELA (–0.698) and math (–0.697) are statistically significant at the 1% and 5% risk levels, respectively, and they are of a similar magnitude as the corresponding baseline estimates.

Adjacent and Older Cohorts. As discussed above, fertility shock could have affected the 2001 cohort, as the Chinese Dragon year ended on January 23, 2001. Furthermore, when there are spillovers across grades within schools, the adjacent cohorts are the most likely to be affected due to having the smallest age difference relative to the Dragon cohort. Thus, we examine the extent to which these potentially contaminated cohorts affect our estimates by excluding them from the sample (column 7 of table 4). This has very little impact on the estimate for math; if anything, the estimate for ELA is higher in this case. This suggests that neither the potential selection effects nor the spillover effects between the Dragon and adjacent cohorts (or the students born in the first 23 days of 2001) significantly affect our results. Column 8 shows the results obtained after balancing the number of cohorts born before and after the Dragon cohort. Excluding the cohorts born before 1997 from the sample also has little effect on the results.

Placebo Tests. To lend further credibility to the validity of our IV strategy, column 9 of table 4 shows the results from a placebo test; this is based on a similar reduced-form specification as the one used in column 2 of table 3, but it excludes the affected Dragon cohort from the sample and replaces the instrument with a pseudo-instrument. We construct this pseudo-instrument with an interaction between the 1990 local share of the Chinese population and a binary indicator that is equal to 1 for the cohorts that are one to three years older than the Dragon cohort (and 0 otherwise). The coefficient of this interaction term indicates whether the non-Asian cohorts that preceded the Dragon cohort performed differently, compared to other non-Asian cohorts within the same school, in areas where the historical share of the Chinese population is larger. These estimates are small and insignificant. We obtain similar findings when we exclude the two potentially contaminated adjacent cohorts from the sample (column 10). Overall, these placebo tests provide further reassurance that the negative effects that an increased share of Asian students has on non-Asian test scores are the result of a plausibly exogenous increase in the share of Asian students in the Dragon cohort stemming from a larger fertility shock in the Dragon year in areas with high historical local shares of the Chinese population.[26]

4.3 Heterogeneity by Race/Ethnicity and Grade

In what follows, we examine whether the negative impacts on non-Asian test scores arise among specific ethnic/racial groups or in certain grades.

Effects by Race/Ethnicity. Panel A of table 5 shows the IV results for different racial/ethnic subgroups. These IV estimates indicate a significant negative effect on the ELA scores of Hispanic students. The point estimate for black students is relatively large and negative for ELA, but its precision is low due to the weaker first stage in this sample (see appendix table A8). In terms of math scores, all groups have negative point estimates, but the coefficient is significant only for the Hispanic group. The only positive (but insignificant) coefficient is for the ELA score of white students, whose achievement distribution is very similar than that of the Asian students in ELA. Although caution is in order when interpreting these results due to the relatively low precision of some estimates, the coefficients do indicate fairly similar negative effects on math scores across all non-Asian ethnic/racial subgroups; the negative impact on ELA scores seems to be mainly the result of adverse effects on Hispanic and black students.

Effects by Grade. Panel B reports IV estimates by grade for third through sixth grade.[27] All the first-stage coefficients on the instrument are larger than 0.11 and are significant at the 5% risk level except for those for sixth grade, which have a smaller sample size (see appendix table A9). The IV estimates suggest that the negative impacts that Asian student share has on non-Asian test scores are fairly similar across grades: The estimates are negative for both subjects across all grades. The largest and most precise point estimates are detected in third and fifth grade.

---

[26] Appendix table A6 shows that IV estimates are only slightly affected when the observations at the right tail of the distributions of initial enrollment and historical Chinese exposure are dropped.

Appendix table A7 shows the first-stage, reduced-form, and IV estimates for the alternative definitions of the instrument. In particular, it shows the results for the specifications of the instrument that are based on various Asian subpopulations (Chinese, Asian excluding Indian, and all Asian) and school neighborhoods (census tracts within 500, 1,000, 2,000, and 3,000 meters of the school). The IV estimates for ELA range from –0.673 to –1.035 across instruments and are all statistically significant ($p<0.05$ for six specifications and $p<0.01$ for the other six). For math, the IV estimates range from –0.358 to –0.639, with seven of the 12 estimates significant at the 5% risk level or lower and two others significant at the 10% risk level. The smallest buffer of 500 meters has the smallest p values. This is likely due to the fact that defining school neighborhoods with a wider radius leads to more overlap between school catchment areas. (We cannot explicitly test for this, however, as we do not know the actual catchment areas.)

[27] Note that sixth grade is the last one in which we observe the Dragon cohort.

4.4 Heterogeneity by School Characteristics

In this section, we examine whether specific school characteristics amplify the impacts of Asian segregation. This analysis could help in identifying the schools that are most negatively affected by the rising share of Asian students. We estimate the following reduced-form regression:

$$y_{rsgt} = \rho_1 CS_{s,1990} + \rho_2 Dragon_{gt} + \tau_1 CS_{s,1990} \times Dragon_{gt} + \tau_2 Z_s \times Dragon_{gt} \\ +\tau_3 CS_{s,1990} \times Z_s + \tau_4 CS_{s,1990} \times Dragon_{gt} \times Z_s + \beta'_1 X_{rsgt} + u_{rsgt}. \quad (3)$$

Here, Z_s is a school characteristic in the academic year 2004/2005 (unless otherwise specified), as that is the first year when school-level information is available in the school-level data.[28] The parameter of interest is the coefficient on the term for the interaction between the school characteristic and the instrument, τ_4. This coefficient tests whether the instrument has a different reduced-form impact with respect to the school-level variable Z_s. The vector X_{rsgt} includes the school, cohort, race, and grade-by-year fixed effects. We weight the regressions by the number of students who took the test and cluster the standard errors at the census-tract level.[29]

We start by examining whether the impacts are heterogeneous by racial/ethnic fractionalization of the non-Asian student population. It could be, for instance, that a larger share of Asian students reduces the size of the accessible peer networks for non-Asian students, which in turn could have a negative impact on those students' average achievement. In this case, the negative effects that result from the entry of additional Asian students could be smaller in schools with less heterogeneous (and possibly more cohesive) non-Asian student populations. We measure fractionalization with the ethno-linguistic fractionalization (ELF) index, $F_s = 1 - \sum_r r_{sr}^2$ (see, e.g., Bossert et al., 2011), where r_{sr} is the share of students in a non-Asian ethnic/racial subgroup r (white, Hispanic, or black) in school s in 2006.[30] This index measures the probability that two randomly selected individuals from the population of non-Asian students would belong to different racial/ethnic groups.[31] Table 6 reports the results. Panel A shows that we detect a

[28] The earlier school-level information is incomplete and unavailable for many of the schools in our sample. See appendix A.2 for more information.

[29] It is worth noting that the inclusion of school fixed effects makes the terms $CS_{s,1990}$ and $CS_{s,1990} * Z_s$ redundant and that the inclusion of cohort fixed effects makes the term $Dragon_{gt}$ redundant.

[30] We use the index of fractionalization from 2006 because racial/ethnic shares were not recorded in the school-level data in 2004/2005. Similarly, we use test scores from 2006 to measure initial school quality.

[31] When all non-Asian students in a school belong to the same racial/ethnic group, this index is equal to 0. The maximum value if there are three non-Asian groups present in a school is $1 - 3 \cdot (1/3)^2 \approx 0.77$.

statistically significant negative coefficient on the interaction term for ELA (p<0.01), which indicates that, given an increase in the Asian population, non-Asians' ELA scores decrease more in schools with more fractionalized non-Asian student populations. For math scores, the corresponding coefficient is also negative but has a smaller magnitude that is not significantly different from 0.

Researchers have shown that larger classes can have negative effects on scholastic achievement (e.g., Angrist and Lavy, 1999; Chetty et al., 2011; Fredriksson et al., 2013). We show below that controlling for average class size has little impact on our estimates, but the impacts of changes in racial composition could interact with class size. Hence, in panel B, we test whether non-Asian students in schools with larger average class sizes are more negatively affected by an increased share of Asian students than those in schools with smaller class sizes, but we find no evidence of this. We also examine whether changes in racial composition interact with school quality (as measured using 2006 test scores), teacher quality, or socioeconomic composition (based on the share of students eligible for free or reduced-price lunches). We do not detect statistically significant differences in the impacts that the share of Asian students has on non-Asian test scores by these variables (see appendix table A10).

5 Alternative Mechanisms

In this section, we assess the importance of some alternative mechanisms (e.g., changes in teaching resources and student attrition) that could explain the negative effects we have identified. We also examine the impacts on the shares of non-Asian students in four proficiency groups to assess how the increase in the share of Asian students affects each part of the non-Asian achievement distribution.

5.1 School-Level Responses

We next examine whether the entry of additional Asian students generates school-level responses. These responses could affect the scholastic achievement of non-Asian students and thus could explain our results. For this purpose, we use annual outcomes from the school-level data. We estimate the following school-level, reduced-form regression:

$$y_{st} = \alpha_s + \alpha_t + \rho_3 Dragon\_In\_School_{st} + \rho_4 Dragon\_In\_School_{st} \times CS_{s,1990} + \epsilon_{st} \quad (4)$$

for school s in year t. Here, $Dragon\_In\_School_{st}$ is a dummy variable that is equal to 1 when the Dragon cohort is in school s and equal to 0 otherwise; y_{st} is a school-level outcome that is drawn from the School Report Cards.[32] The coefficient of the interaction term ρ_4 detects the differences in the impact of the Dragon cohort across schools in areas with low and high Chinese population shares in 1990. We include school and year fixed effects, and we cluster the standard errors at the census-tract level.

School-Level Enrollment. Table 7 shows the results. We detect no statistically significant differential impacts of the Dragon cohort on school-level enrollment between areas with high and low historical Chinese population shares. The small positive point estimate for enrollment at the school-year level is in line with the positive effect that the instrument had on enrollment at the school-grade-year level (table 2, column 6). Arguably, the impact of the disproportionately large Dragon cohort on enrollment is significantly attenuated at the school level.

Class Size and Teacher Resources. We showed above that the instrument has no impact on grade-level average class size. As a complementary analysis, we run the school-level regression (4) by using annual school-level average class size from grades three through six as the outcome. We also examine the impact of the entry of the Dragon cohort on the number of teachers and students per teacher ratio. We find that the point estimates for these outcomes are all insignificant and small, which suggests that class size and teacher resources were not significantly different in areas that had higher and lower historical Chinese population shares when the Dragon cohort entered the school. A potential explanation for these findings is that primary-school resource allocation is based on student count, which would mean that additional students would not have significant impacts on teachers' resources per student. Moreover, class-size rules prevent the average class size from increasing when additional students are enrolled in a school.

Teacher Quality. The positive (but small and insignificant) estimate in column 2 of table 7 suggests that some schools might have recruited new teachers in response to the enrollment of the additional Asian students. This raises the concern that new recruitment affected teacher quality, which in turn may have affected test scores. We investigate this possibility by using (as dependent variables) school-level measures of teachers' qualifications and experience. Using these specifications, we do not detect statistically significant impacts on the share of teachers who lack a valid certificate. The impact on the

---

[32] To construct the $Dragon\_In\_School_{st}$ dummy, we employ information on the minimum and maximum grades in each school (see appendix table A1). Because we are interested in a school's response to the entry of the Dragon cohort we exclude school-year observations from after the Dragon cohort exited the school. For example, a school that starts in kindergarten and continues through fifth grade has a dummy of 0 in 2004/2005, a dummy of 1 from 2005/2006 (when the Dragon cohort attends kindergarten) until 2010/2011 (when the Dragon cohort attends the fifth grade).

percentage of teachers with fewer than three years of experience is also insignificant. These findings provide little support for the hypothesis that reductions in teacher quality drove the adverse impacts on non-Asian scholastic achievement.

Socioeconomic Status, Disruption, and Attendance. Columns 10 and 11 of table 7 show the results for the percentage of students who are eligible for free or reduced-price lunch. Again, the impacts on these outcomes are small, suggesting that the percentage of low-income students did not increase when the Dragon cohort entered the school. Negative impacts on test scores could also result from increased disruption (e.g., Lazear, 2001; Hanushek et al., 2004). To investigate this hypothesis, we use the last two columns to examine disruption's potential effects on the share of suspended students and on attendance rates. We do not detect impacts for these outcomes, suggesting that the entry of additional Asian students does not induce changes in absence or disruptive behavior.

School-Level Variables as Added Covariates. We now add school-level variables as covariates in our IV model. This is an alternative strategy for assessing whether a change in the share of Asian students affects non-Asian scholastic performance through school-level responses. Panel A in appendix table A11 shows the results for 912 schools for which School Report Card variables are available. For this sample, the estimates corresponding to our baseline specification (–0.781 with $p<0.01$ for ELA; –0.622 with $p<0.05$ for math) are very similar to the corresponding estimates in the full sample. These estimates are slightly smaller than in the full sample (but are still large in magnitude and significant at the 5% risk level) when we include the full set of school-level variables, as well as the shares of white and Hispanic students, to control for changes in the racial/ethnic composition of non-Asian student population at the grade level (–0.632 for ELA; –0.547 for math).[33] In panel B, we show similar specifications when grade-level average class size is added in the set of control variables. This variable is unavailable for the academic year 2005/2006, which reduces the sample size further. The estimates for the baseline specification (–0.594 with $p<0.05$ for ELA; –0.569 with $p<0.05$ for math) are to some extent smaller compared to the corresponding estimates in the full sample. Controlling for the full set of control variables, including the grade-level average class size, reduces the estimates slightly, but they are both economically and statistically significant (–0.495 with $p<0.1$ for ELA; –0.503 with $p<0.05$ for math).

[33] Note that the share of black students is omitted because the shares sum to unity, making the black-student share redundant when the shares of Asian, Hispanic, and white students are included in the regression.

The evidence presented above suggests that school-level responses to the enrollment of additional Asian students are unlikely to have driven our results. We acknowledge, however, that caution is in order when interpreting some of these results. For instance, we observe teachers' qualifications and experience at the school level, but we cannot rule out the possibility that these factors vary across grades within schools. However, if principals aim to allocate resources evenly across grades and to increase resources for any grades for which student performance is declining, our negative estimates of the effects on non-Asian achievement are likely to be biased toward zero (e.g., Ballatore et al., forthcoming).

5.2 Attrition

In this section, we examine the extent to which our estimates could have been affected by the potential attrition of non-Asian students. For instance, the parents of some negatively affected non-Asian students may have decided to transfer their children to another school because of weak academic performance. This, in turn, could have shifted the non-Asian achievement distribution, inducing confounding variation in non-Asian test scores. To test for endogenous attrition, we track cohorts across grades within schools and examine the impact that the instrument has on the changes in the student populations between grades. Table 8 shows the results for changes between third and fourth grade and between fourth and fifth grade for white, black, and Hispanic students.[34] The instrument's impact on the change in the number of students from third to fourth grade is small and insignificant for each group. The impact on the change from fourth to fifth grade is small and insignificant for black and white students and *positive* and significant for Hispanic students (0.03, $p<0.05$). The magnitude of the latter estimate is relatively small, given that the dependent variable mean is 46.8 Hispanic students per school-grade-year cell. Moreover, the positive point estimate does not support the hypothesis that negative test-score impacts induce endogenous attrition among students. This is consistent with the fact that school transfers are granted only in rare circumstances and suggests that the attrition of non-Asian students is unlikely to have driven our estimates.

[34] We restrict this analysis to changes across these grades because, in most primary schools, fifth grade is the last grade (see appendix table A1) and therefore the number of observations is limited for examining attrition from fifth grade onward.

5.3 Effects on Proficiency

In an attempt to further understand the channels that generated the negative impacts on non-Asian performance, we examine the effects that the share of Asian students has on the non-Asian proficiency distribution. Table 9 shows the results for IV specifications, using the share of non-Asian students in four proficiency groups as the dependent variable. Panels A, B, and C show the results for all non-Asian students, for white students, and for Hispanic and black students, respectively. For these specifications, the instrument's first-stage coefficients are highly significant (see appendix table A12).

For the sample that includes all non-Asian students, the IV estimate for math shows a statistically significant positive effect on the share of students in the lowest proficiency group (0.54; $p<0.05$). For ELA, however, we do not detect a similar impact on the lowest-performing group, but we do see a positive impact on the share of students in the second-lowest proficiency group, which includes students who are below proficiency (0.45; $p<0.10$). Looking at the two highest proficiency groups, we detect similar negative impacts on the share of students in the highest ELA group (–0.332; $p<0.05$) and in the second-highest math group (–0.367; $p<0.10$). A comparison of the pattern of the estimates for the two subject areas suggests that the significant coefficients for math are at one proficiency level lower than they are for ELA. In particular, the sharp rise in the share of students in the lowest math proficiency group suggests that the additional Asian students caused a larger fraction of non-Asian students to lag behind in this subject. The magnitude of this effect is also large in relative terms, as the average share of students in the lowest math proficiency group is relatively small: around 7.7%. For ELA, the pattern of estimates is not as dramatic in the lower part of the proficiency distribution. However, the results suggest that a major factor in the negative impacts on average non-Asian ELA achievement is the reduction in the share of high-performing students.[35]

5.4 Discussion of Potential Mechanisms

The sharp increase in the share of students in the poorest-performing math proficiency group (as shown in table 9) could be a result of an increase in the pace and coverage of

[35] The pattern of the effects (shown in panel C) for Hispanic and black students is very similar in the two subjects. For white students, we detect a negative effect on the share of students with the highest math proficiency (–0.49), a negative impact on the share of students well below proficiency (–0.14), and a positive impact on the share of middle-level proficiencies. However, these coefficients are not significantly different from 0. The effects on the shares of white students in the ELA proficiency groups are neither sizeable nor significant.

instruction due to the larger share of Asian students, who are particularly high-achieving in math (see appendix figure A2). This would be expected to negatively affect the students in the lower parts of the proficiency distribution and could result in an increasing proportion of them lagging behind. The estimated impacts are less dramatic in the lower part of the ELA proficiency distribution, however. For this subject, we observe a leftward shift in the proficiency distribution combined with a compression of the distribution. This pattern is consistent with a reduced pace and level of ELA instruction, as a slower pace could help prevent less-well-performing students from falling well below proficiency; however, this could also reduce the share of students who perform extremely well.[36]

Other mechanisms could explain our findings as well. In our data, a typical additional Asian student resulting from the fertility spike in the Dragon year is, on average, high-achieving, as are the Asian students in other cohorts. Therefore, our findings could be explained by negative peer effects that arise from having high-achieving peers.[37] It might be, for instance, that well-performing Asian students discourage non-Asian students, thus reducing their study effort.[38] Negative peer influences could also result from changes in average peer personality.[39] Finally, the importance of these mechanisms could depend on how parents react to changes in the school environment and in the teaching of their children (e.g., Pop-Eleches and Urquiola, 2013; Fredriksson et al., 2016).

[36] One potential explanation for the possible reduction in the pace of ELA instruction is that a relatively large share of Asian children have limited English proficiency, so an increase in the share of Asian students could increase the share of all students with limited English proficiency. Although Asian students performed relatively well on the ELA test, 27% of those who spoke Chinese at home had limited proficiency in English (Asian American Foundation, 2014). The rates of limited English proficiency among children are also high for other large Asian-language groups (e.g., 30% for Vietnamese and 26% for Korean). The negative impacts on test scores of changing the pace of instruction—an increased pace in math and a decreased pace in ELA—can be reconciled using a simple model in which any deviations from the optimal pace of instruction in either direction reduce students' average achievement. This concave relationship between teaching pace and average achievement can emerge when teaching pace has nonlinear effects across the proficiency distribution. This would be the case, for instance, if the benefit to good students is less than the harm to poor students when the teacher goes faster than the optimal pace, or if the harm to good students is greater than the benefit to poor students when the teacher goes slower than the optimal pace.

[37] Many previous studies have suggested that the effects of high-achieving students are heterogeneous across the achievement distribution, but the findings are mixed. Some find that high-achieving students benefit from having high-achieving peers (e.g., Hoxby and Weingarth; 2005; Gibbons and Telhaj, 2008). However, Lefgren (2004), when examining the impacts of tracking based on prior achievement, finds little evidence that high-achieving students benefit from being placed in classes with high-achieving peers. Lavy, Silva, and Weinhardt (2009) find heterogeneous effects along the gender dimension: Girls, especially those in the bottom half of the ability distribution, benefited significantly from interactions with very high-achieving peers. In contrast, boys were marginally negatively affected by having a larger proportion of academically outstanding peers.

[38] For instance, researchers have shown that having knowledge of one's relatively low position in an achievement ranking can be discouraging (e.g., Murphy and Weinhardt, 2016).

[39] See, for example, Golsteyn et al. (2017), who provide evidence that reducing average peer persistence and risk tolerance can have negative impacts on achievement. In the context of our study, it may be, for instance, that an increase in the share of Asian students reduces the average persistence or risk tolerance in the classroom, which could explain the negative effects that we find.

6 Conclusions

The U.S. Asian population grew by 72% between 2000 and 2015, from 11.9 million to 20.4 million, and according to a U.S. Census Bureau projection, this population will account for around 30% of the U.S. immigrant population by 2060 (Colby and Ortman, 2015). The rapidly rising share of the Asian population will likely increase Asian segregation in many parts of the country, but its economic and social consequences are still largely unknown.

In this study, we examine one particularly important aspect of Asian segregation: the extent to which it affects scholastic performance among primary-school students. In our analysis, we employ test-score data for public primary schools in NYC, which houses one of the largest Asian populations in the U.S. We address endogeneity concerns by implementing a novel research design that exploits plausibly exogenous variations in the share of Asian students across schools as a result of the interaction between the fertility spike among the Asian population in the Chinese year of the Dragon and the historical local Chinese population share. The applicability of this empirical strategy is not specific to the schooling context and could also be helpful for future studies that examine other economic and social consequences of racial segregation.

We shed light on this important but difficult question by providing several new results. We find that Asian segregation has a negative and statistically significant causal effect on both ELA and math scores of non-Asian students. The main driver of the negative effect on the ELA score is the adverse effect on Hispanic and black students. Conversely, we find little evidence of the negative effects on the performance of white students. Moreover, effects on math scores are driven by a sharp increase in the share of students who lag behind, which could be a result of an increase in the pace and coverage of instruction due to the larger share of Asian students, who are particularly well-achieving in math. Our findings suggest that reducing Asian segregation can be especially beneficial for minority groups who are, on average, less-well performing. They also provide suggestive evidence that desegregation efforts may generate net benefits in terms of average student achievement.

References:

Adams, John S., William C. Block, Mark Lindberg, Robert McMaster, Steven Ruggles, and Wendy Thomas. *National Historical Geographic Information System: Pre-release Version 0.1*. Minnesota Population Center University of Minnesota, Minneapolis (2004).

Angrist, Joshua D., and Kevin Lang. "Does school integration generate peer effects? Evidence from Boston's Metco Program." *American Economic Review* 94, no. 5 (2004): 1613-1634.

Angrist, Joshua D., and Victor Lavy. "Using Maimonides' rule to estimate the effect of class size on scholastic achievement." *Quarterly Journal of Economics* 114, no. 2 (1999): 533-575.

Asian American Federation. *The State of Asian American Children*. New York, NY (2014).

Ballatore, Rosario M., Margherita Fort, and Andrea Ichino. "The Tower of Babel in the classroom: immigrants and natives in Italian schools." *Journal of Labor Economics* (forthcoming).

Black, Sandra "Do better schools matter? Parental valuation of elementary education." *Quarterly Journal of Economics* 114, no. 2 (1999): 577-599.

Black, Sandra, and Stephen Machin. "Housing valuations of school performance." *Handbook of the Economics of Education* 3 (2011): 485-519.

Bossert, Walter, Conchita D'Ambrosio, and Eliana La Ferrara "A Generalized Index of Fractionalization." *Economica* 78, no. 312 (2011): 723-750.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. "How does your kindergarten classroom affect your earnings? Evidence from Project STAR." *Quarterly Journal of Economics* 126, no. 4 (2011): 1593-1660.

Colby, Sandra L., and Jennifer M. Ortman. "Projections of the size and composition of the U.S. population: 2014 to 2060: Population estimates and projections." U.S. Census Bureau Current Population Reports (2017).

Cook, Michael D., and William N. Evans. "Families or schools? Explaining the convergence in white and black academic performance." *Journal of Labor Economics* 18, no. 4 (2000): 729-754.

Duflo, Esther, Pascaline Dupas, and Michael Kremera. "Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya." *American Economic Review* 101, no. 5 (2011): 1739-1774.

Fredriksson, Peter, Björn Öckert, and Hessel Oosterbeek. "Long-term effects of class size." *Quarterly Journal of Economics* 128, no. 1 (2013): 249-285.

Fredriksson, Peter, Björn Öckert, and Hessel Oosterbeek. "Parental responses to public investments in children: Evidence from a maximum class size rule. *Journal of Human Resources* 51, no. 4(2016), 832-868.

Geay, Charlotte, Sandra McNally, and Shqiponja Telhaj. "Non-native speakers of English in the classroom: What are the effects on pupil performance?" *The Economic Journal* 123, no. 570 (2013): 281-307.

Gibbons, Steve, and Shqiponja Telhaj. *Peers and Achievement in England's Secondary Schools*. Spatial Economics Research Centre Discussion Paper no. 1 (2008).

Golsteyn, Bart, Arjan Non, and Ulf Zölitz. *The impact of peer personality on academic achievement.* University of Zurich, Department of Economics, Working Paper No. 269 (2017).

Goodkind, Daniel M. "The significance of demographic triviality: minority status and zodiacal fertility timing among Chinese Malaysians." *Population Studies* 49, no. 1 (1995): 45-55.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. "Disruption versus Tiebout improvement: the costs and benefits of switching schools." *Journal of Public Economics* 88, no. 9 (2004): 1721-1746.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. "New evidence about *Brown v. Board of Education*: the complex effects of school racial composition on achievement." *Journal of Labor Economics* 27, no. 3 (2009): 349-383.

Hanushek, Eric A., and Steven G. Rivkin. "Harming the best: how schools affect the black-white achievement gap." *Journal of Policy Analysis and Management* 28, no. 3 (2009): 366-393.

Hesketh, Therese, Yan Zhen, Li Lu, Zhou Xu Dong, Ye Xu Jun, and Zhu Wei Xing. "Stress and psychosomatic symptoms in Chinese school children: cross-sectional survey." *Archives of Disease in Childhood* 95, no. 2 (2010): 136-140.

Hoxby, Caroline M. *Peer effects in the classroom: Learning from gender and race variation*. NBER Working Paper no. 7867 (2000).

Hoxby, Caroline M., and Gretchen Weingarth. *Taking race out of the equation: School reassignment and the structure of peer effects*. NBER Conference Paper (2005).

Imberman, Scott A., Adriana D. Kugler, and Bruce I. Sacerdote. "Katrina's children: evidence on the structure of peer effects from hurricane evacuees." *American Economic Review* (2012): 2048-2082.

Jensen, Eric B., Anthony Knapp, Peter Borsella, and Kathleen Nestor. *The Place-of-Birth Composition of Immigrants to the United States: 2000 to 2013.* U.S. Census Bureau Population Division (2015).

Johnson, Noel D., and John V. C. Nye. "Does fortune favor dragons?" *Journal of Economic Behavior & Organization* 78, no. 1 (2011): 85-97.

Lavy, Victor, Olmo Silva, and Felix Weinhardt. "The good, the bad, and the average: evidence on ability peer effects in schools." *Journal of Labor Economics* 30, no.2 (2012): 367-414.

Lavy, Victor, M. Daniele Paserman, and Analia Schlosser. "Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom." *Economic Journal* 122, no. 559 (2012): 208-237.

Lazear, Edward P. "Educational production." *Quarterly Journal of Economics* 116, no. 3 (2001): 777-803.

Lefgren, Lars. "Educational peer effects and the Chicago public schools." *Journal of Urban Economics* 56, no. 2 (2004): 169-191.

Mocan, Naci H., and Han Yu. *Can superstition create a self-fulfilling prophecy? School outcomes of Dragon children of China*. NBER Working Paper no. 23709 (2017).

Murphy, Richard, and Felix Weinhardt. *Top of the class: The importance of ordinal rank*. CESifo Working Paper No. 4815 (2014).

Pop-Eleches, Cristian, and Miguel Urquiola. "Going to a Better School: Effects and Behavioral Responses." *American Economic Review* 103, no. 4 (2013): 1289–1324.

Rivkin, Steven, and Finis Welch. "Has school desegregation improved academic and economic outcomes for blacks?" In *Handbook of the Economics of Education* 2, edited by Eric A. Hanushek and Finis Welch (2006): 1019-1049.

Sacerdote, Bruce. "Peer effects in education: How might they work, how big are they and how much do we know thus far?" *Handbook of the Economics of Education* 3, no. 3 (2011): 249-277.

Yip, Paul S. F., Joseph Lee, and Yin-Bun Cheung. "The influence of the Chinese zodiac on fertility in Hong Kong SAR." *Social Science & Medicine* 55, no. 10 (2002): 1803-1812.

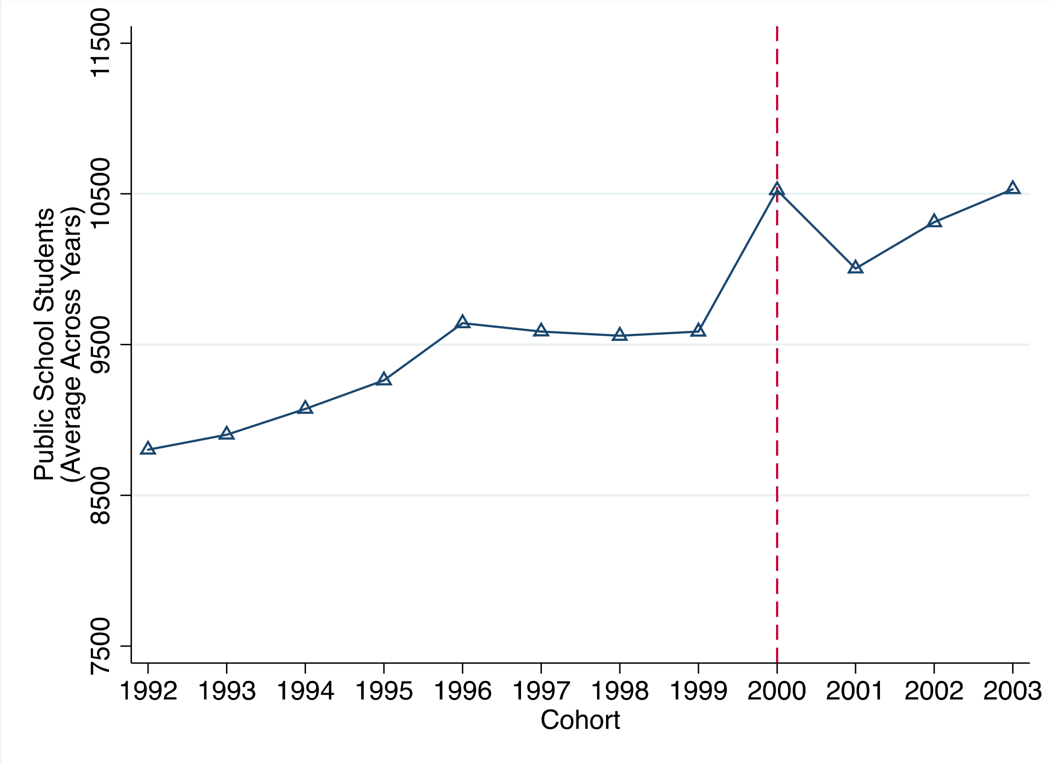

Figure 1: Number of Asian Students in New York City Public Schools by Birth Cohort

Notes: This figure displays the average number of Asian primary public-school students by birth cohort. It is based on data on the number of students completing the ELA or math test by school, year, grade, and ethnicity/race. If the number of students taking the test is not the same between the two tests within a cell, the larger value is used. The figures are constructed by aggregating the number of students into the year-cohort level and averaging the year-cohort figures by cohort. Data source: New York City Department of Education mean test score files by school, grade, year, and ethnicity/race for the years of 2006-2012.

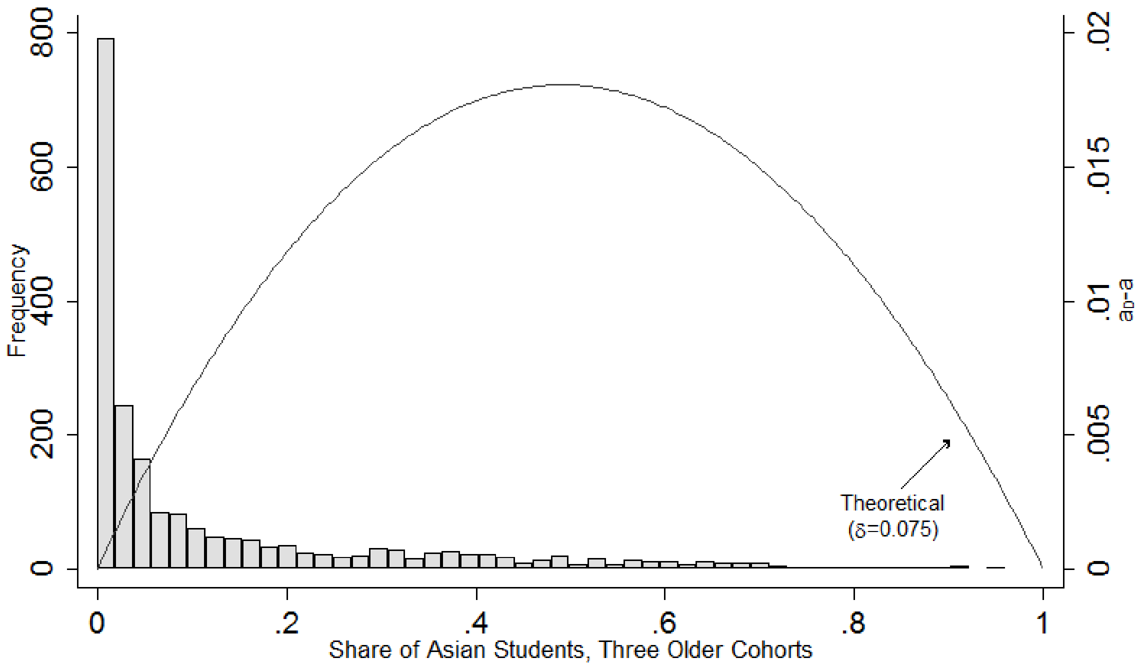

Figure 2: Theoretical Impact of the Fertility Shock in the Dragon Year

Notes: This figure displays the theoretical impact of a fertility shock increasing the number of Asian children between an initial and affected birth cohort by 7.5%, keeping the number of non-Asian children across cohorts fixed, as a function of the share of Asian children in the initial birth cohort (right axis). The histogram shows the empirical distribution of the share of Asian students in third grade for three pre-Dragon birth cohorts born in 1997-1999 (left axis).

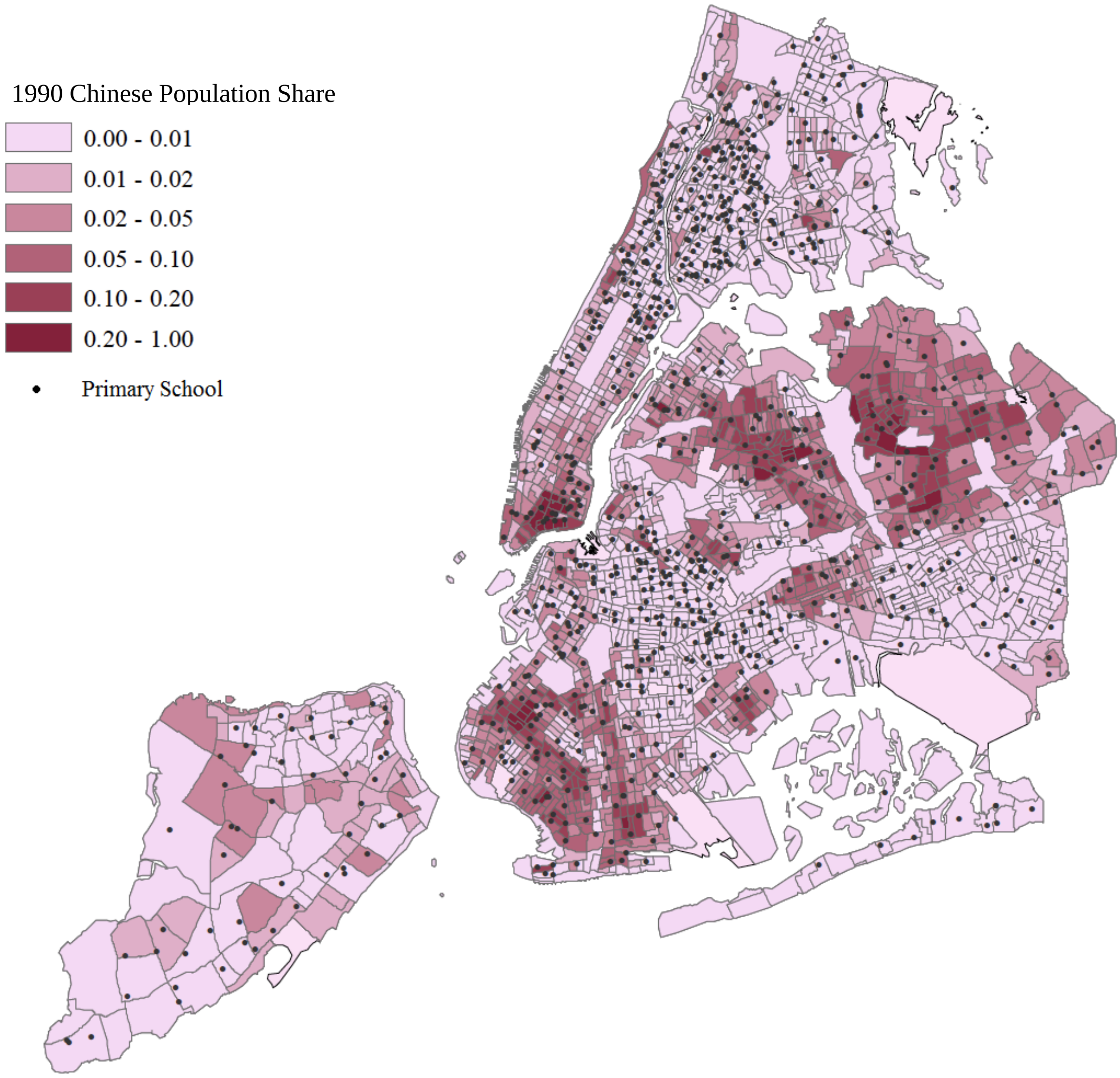

Figure 3: Chinese Population Share by Census Tract in New York City, 1990

Notes: This figure displays the Chinese population share by census tract in New York City in 1990. Dots represent schools in our data. Data sources: U.S. Census Bureau 1990 census data by census tract and 1990 census-tract shapefiles, both provided by the Minnesota Population Center; New York State School Report Cards.

Table 1
Summary Statistics

	Full Sample			1990 Chinese Population Share: Below Median			1990 Chinese Population Share: Above Median		
	Mean	Standard Deviation	N	Mean	Standard Deviation	N	Mean	Standard Deviation	N
Panel A. Test Scores									
Math, All	675.1	(23.0)	59,399	667.9	(20.9)	29,789	682.9	(22.6)	29,610
Math, Asian	701.2	(17.1)	9,527	695.0	(19.8)	2,533	702.2	(16.4)	6,994
Math, White	692.4	(17.8)	9,141	691.4	(17.7)	2,746	693.0	(17.8)	6,395
Math, Black	664.2	(18.1)	18,824	663.1	(18.1)	12,113	667.1	(17.9)	6,711
Math, Hispanic	667.5	(18.3)	21,907	664.4	(18.4)	12,397	671.6	(17.4)	9,510
Math, Non-Asian	670.6	(20.7)	49,872	666.7	(20.1)	27,256	676.1	(20.4)	22,616
ELA, All	658.6	(16.5)	59,399	654.0	(15.3)	29,789	663.6	(16.3)	29,610
ELA, Asian	672.4	(14.5)	9,527	669.9	(14.8)	2,533	672.8	(14.4)	6,994
ELA, White	674.1	(14.6)	9,141	673.7	(13.9)	2,746	674.4	(15.0)	6,395
ELA, Black	653.0	(12.7)	18,824	652.0	(12.5)	12,113	655.6	(13.0)	6,711
ELA, Hispanic	652.1	(13.3)	21,907	649.6	(13.4)	12,397	655.3	(12.6)	9,510
ELA, Non-Asian	656.2	(15.6)	49,872	653.3	(14.9)	27,256	660.3	(15.7)	22,616
Panel B. Enrollment, Student Shares, and Class Size									
Enrollment	125.1	101.4	17,329	110.4	80.3	9,408	142.6	119.5	7,921
Asian-Student Share	13.3	18.1	17,329	4.3	8.6	9,408	24.0	20.6	7,921
Black-Student Share	32.2	28.9	17,329	43.0	29.4	9,408	19.3	22.2	7,921
Hispanic-Student Share	39.6	25.7	17,329	43.4	27.2	9,408	35.1	23.0	7,921
White-Student Share	14.9	21.2	17,329	9.3	20.1	9,408	21.6	20.5	7,921
Class Size, Grade-Level Average	24.8	4.5	14,464	24.2	4.5	7,890	25.4	4.5	6,574

Notes: The data include 59,399 observations at the school-year-grade-ethnicity/race level for 1,081 New York City public primary schools, and they cover the years 2006-2012 and include third through eighth grade. Descriptive statistics are presented for the full sample and for schools below and above the median of the 1990 Chinese population share, which is calculated by using census tracts within 500 meters of the school. Data on ELA and math scores represent averages at the school-grade-year-race/ethnicity level. Sample means for these variables are weighted by the number of students attending the test (panel A). Enrollment is measured by the number of students attending the tests (when attendance in the two tests is not equal, the maximum value is used). Student shares by race/ethnicity are based on enrollment figures at the school-grade-year-ethnicity/race level. Grade-level average class size is observed for a subsample excluding the academic year 2005/2006. Sample means in panel B are unweighted.

Table 2

Impact of the Instrument on Enrollment and Asian Test Scores in Third Grade

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Enrollment, Third Grade						Asian Test Scores, Third Grade		Average Grade Size, Third Grade
	Asian	Non-Asian	White	Black	Hispanic	All	Math	ELA	
% Chinese 1990 × Dragon	0.1230**	–0.0136	0.0233	–0.0119	-0.0632	0.1090**	0.0250	0.0205	-0.0131
	(0.0517)	(0.0363)	(0.0267)	(0.0364)	(0.0447)	(0.0507)	(0.0440)	(0.0410)	(0.0179)
Outcome Mean	13.9	84.4	14.5	27.8	39.4	98.3	704.9	676.2	22.43
Observations	4,917	4,917	4,917	4,917	4,917	4,917	1,971	1,971	2,898
Number of Schools	722	722	722	722	722	722	328	328	680

Notes: Columns 1-6 report reduced-form estimates of the effect of the instrument on enrollment in third grade. Columns 7-8 and 9 report reduced-form effects on Asian test scores and grade-level average class size in third grade, respectively. The specifications in columns 7-8 have fewer observations because all schools do not have Asian students who undertake the tests. The specification in column 9 has fewer observations because the outcome is unavailable for the academic year 2005/2006. The unit of observation is at the school-cohort-year level. The instrument is the interaction between the Dragon dummy, equal to 1 for the Dragon cohort and 0 otherwise, and the 1990 Chinese population share in census tracts within 500 meters of the school. All specifications include school, cohort, and year fixed effects. Columns 2 and 6 also include race fixed effects. Standard errors (in parentheses) are clustered at the census-tract level. *** p<0.01, ** p<0.05, * p<0.10.

Table 3

Effects of the Share of Asian Students on Non-Asian ELA and Math Scores

	(1) First Stage	(2) Reduced Form	(3) IV	(4) OLS
Panel A. Outcome: Non-Asian ELA Score				
% Chinese 1990 × Dragon	0.151**	–0.116***		
	(0.0600)	(0.0447)		
Asian Share			–0.765**	0.00489
			(0.314)	(0.0461)
Panel B. Outcome: Non-Asian Math Score				
% Chinese 1990 × Dragon	0.153**	–0.0979**		
	(0.0603)	(0.0438)		
Asian Share			–0.639**	0.0521
			(0.311)	(0.0406)
Observations	49,872	49,872	49,872	49,872
Number of Schools	1,080	1,080	1,080	1,080

Notes: This table reports coefficients from IV and OLS regressions of ELA and math scores of non-Asian students on the share of Asian students. Outcomes are mean test scores by school-year-grade-ethnicity/race cells. The instrument is the interaction term between the Dragon cohort dummy and the local 1990 Chinese population share. All specifications control for school, cohort, year, grade, grade-by-year, and race fixed effects. Columns 1 and 2 display the first-stage and reduced-form coefficients on the instrument. Columns 3 and 4 display the IV and OLS coefficients on the share of Asian students. Regressions are weighted by the number of students completing the test. Standard errors (in parentheses) are clustered at the census-tract level. *** p<0.01, ** p<0.05, * p<0.10.

Table 4

Robustness Analysis: Local Trends, Adjacent Cohorts, and Placebo Tests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	IV								Reduced-Form	
	Baseline	Quadratic Instrument	Quadratic Year Trend × 1990 Chinese Share	Column 3 + Quadratic Birth Year Trend × 1990 Chinese Share	Column 4 + Quadratic Instrument	Column 5 + Controls for Grade Size	No Adjacent Cohorts	No Old Cohorts	Placebo: Three Older Cohorts	Placebo: Column 9 Excluding Adjacent Cohorts
ELA Non-Asian	–0.765**	–0.813**	–0.597*	–0.570	–0.801*	–0.698***	–0.955***	–0.731**	0.002	–0.006
	(0.314)	(0.349)	(0.307)	(0.354)	(0.449)	(0.265)	(0.351)	(0.341)	(0.031)	(0.036)
Math Non-Asian	–0.639**	–0.660**	–0.949**	–0.796*	–0.760*	–0.697**	–0.695**	–0.695*	0.019	0.021
	(0.311)	(0.295)	(0.415)	(0.479)	(0.427)	(0.301)	(0.321)	(0.361)	(0.40)	(0.039)
Observations	49,872	49,872	49,872	49,872	49,872	49,872	39,230	34,838	44,546	33,904
Number of Schools	1,080	1,080	1,080	1,080	1,080	1,080	1,080	1,080	1,080	1,080

Notes: This table reports IV estimates of the effect of the Asian-student share on non-Asian ELA and math scores (columns 1–8) and reduced-form placebo tests (columns 9-10). All specifications control for school, cohort, grade-by-year, and race fixed effects. Column 1 displays the baseline estimates corresponding to the IV estimates in Table 3. Column 2 adds the square of the local 1990 Chinese population share interacted with the Dragon dummy as an instrument. Column 3 is similar to column 1 but controls for a second-order polynomial of year interacted with the local 1990 Chinese population share. Column 4 adds a second-order polynomial of birth year interacted with the local 1990 Chinese population share. Column 5 is similar to column 4 but adds the square of the instrument as an additional instrument. Column 6 adds control variables for grade size and its square, and it adds 10 dummies for grade-size deciles. Column 7 excludes cohorts born one year before and after the Dragon cohort. Column 8 excludes cohorts born before 1997. Column 9 displays reduced-form placebo tests for regressions that correspond to the specification in column 2 in Table 3 but that replace the instrument with the interaction of the 1990 local Chinese population share and a dummy variable equal to 1 for birth cohorts from 1997 to 1999 and 0 otherwise; the regressions also exclude the Dragon cohort from the sample. The specification in column 10 is the same as in column 9 but excludes cohorts adjacent to the Dragon cohort from the sample. Regressions are weighted by the number of students attending the test. Standard errors (in parentheses) are clustered at the census-tract level. *** p<0.01, ** p<0.05, * p<0.10.

Table 5
Effects by Race/Ethnicity and Grade

	(1)	(2)	(3)	(4)
Panel A: Race/Ethnicity	White	Hispanic	Black	Hispanic and Black
ELA	0.377	–0.591*	–1.409	–0.571*
	(0.454)	(0.322)	(0.992)	(0.295)
Math	–0.451	–0.560*	–0.884	–0.549
	(0.399)	(0.329)	(1.126)	(0.350)
N	9,104	21,907	18,824	40,731
Panel B: Grade	3rd	4th	5th	6th
ELA	–0.741*	–0.386	–1.149**	–0.596
	(0.406)	(0.460)	(0.540)	(0.572)
Math	–1.240**	–0.396	–0.764*	–0.484
	(0.532)	(0.379)	(0.434)	(0.754)
N	9,982	9,937	9,930	7,216

Notes: This table reports IV estimates of the effect of the share of Asian students on non-Asian ELA and math test scores by race/ethnicity and grade. All specifications include school, cohort, year, grade, grade-by-year (except panel B), and race (except panel A, columns 1–3) fixed effects. Regressions are weighted by the number of students completing the test. Standard errors (in parentheses) are clustered at the census-tract level. *** p<0.01, ** p<0.05, * p<0.10.

Table 6
Heterogeneity by School Characteristics

	(1) ELA	(2) Math
Panel A. Fractionalization Index × % Chinese 1990 × Dragon	–0.628***	–0.200
N=47,663	(0.235)	(0.213)
Panel B. Class Size × % Chinese 1990 × Dragon	0.00762	–0.00350
N=39,232	(0.0170)	(0.00826)

Notes: This table reports reduced-from coefficients on the interaction between the instrument and a school characteristic. The index of fractionalization is measured in 2005/2006, and class size in 2004/2005. In both cases, these are the first years when information on these variables is available in the data. Each cell reports a coefficient from a separate regression. For example, the estimate in the first row of column 1 is the coefficient on the interaction between the 1990 Chinese population share, Dragon dummy, and index of fractionalization for ELA. All specifications control for the main effects and the interaction between the Dragon dummy and school characteristic, the interaction between the Dragon dummy and 1990 Chinese population share, and the interaction between the school characteristic and 1990 Chinese population share, and they include school, cohort, grade-by-year, and race fixed effects. The lower number of observations for class size in panel B is due to the unavailability of data for some schools. Regressions are weighted by the number of students completing the test. Standard errors (in parentheses) are clustered at the census-tract level. *** p<0.01, ** p<0.05, * p<0.10.

Table 7
School-Level Responses

Panel A

Dependent Variable:	(1) Student Enrollment	(2) Number of Teachers	(3) Students per Teacher	(4) Average Class Size	(5) Number of Classes	(6) % Teachers without Valid Certification	(7) % Teachers Out of Certification
Dragon in School × % Chinese 1990	0.143	0.0021	–0.0137	0.0453	–0.110	0.0637	0.0317
	(1.258)	(0.0477)	(0.0167)	(0.0656)	(0.226)	(0.0694)	(0.0672)
Outcome Mean	629.5	48.2	13	24.4	103	3	9.6

Panel B

Dependent Variable:	(8) % Teachers with Less Than 3 Years of Experience	(9) % Teachers without Appropriate Certification	(10) % Students with Free-Lunch Status	(11) % Students with Reduced-Lunch Status	(12) % Suspended Students	(13) % Attendance
Dragon in School × % Chinese 1990	0.0585	–0.0089	0.0919	0.0081	–0.0288	–0.0511
	(0.142)	(0.0716)	(0.105)	(0.0154)	(0.0395)	(0.0779)
Outcome Mean	14.3	11.8	68.9	8.4	3.4	90.9

Notes: This table displays school-year regressions using school-level outcomes indicated by the column title. The sample includes 7,384 observations in 1,059 schools except for in column 4, where it includes 5,715 observations in 873 schools. Data are for academic years 2004/2005 to 2011/2012. Each regression includes school fixed effects. Dragon in School is a dummy variable indicating whether the Dragon cohort is in a school. The coefficient on its interaction with the 1990 Chinese population share provides a test for whether schools with larger historical Chinese population shares responded to the entry of the Dragon cohort differently compared with schools with lower shares. The lower number of observations for class size in column 4 is due to the unavailability of data for some schools. Standard errors are clustered at the census-tract level. *** p<0.01, ** p<0.05, * p<0.10.

Table 8

Testing for Attrition

	(1)	(2)	(3)	(4)	(5)	(6)
	Change in Enrollment					
	From Grades 3 to 4			From Grades 4 to 5		
	White	Black	Hispanic	White	Black	Hispanic
% Chinese 1990 × Dragon	0.0059	0.014	0.013	–0.018	0.026	0.032**
	(0.0097)	(0.025)	(0.018)	(0.016)	(0.016)	(0.016)
Observations	4,153	4,153	4,153	4,095	4,095	4,095
Number of Schools	707	707	707	696	696	696

Notes: This table reports reduced-form estimates of the effect of the instrument on the change in the number of students from grade g to grade $g + 1$ within school and cohort. All specifications include school, cohort, and year fixed effects. The unit of observation is a school-cohort. Standard errors (in parentheses) are clustered at the census-tract level. *** p<0.01, ** p<0.05, * p<0.10.

Table 9
Effects on Proficiency

	(1)	(2)	(3)	(4)
Proficiency Level:	1 Well Below Proficiency	2 Below Proficiency	3 Proficient	4 Above Proficiency
Panel A: All Non-Asian				
ELA	–0.281	0.454*	0.158	–0.332**
	(0.258)	(0.242)	(0.303)	(0.136)
Math	0.537**	0.111	–0.367*	–0.282
	(0.265)	(0.229)	(0.214)	(0.295)
N=49,872				
Panel B: White				
ELA	–0.0820	–0.0195	–0.0310	0.135
	(0.181)	(0.387)	(0.492)	(0.311)
Math	–0.140	0.380	0.248	–0.490
	(0.134)	(0.348)	(0.381)	(0.523)
N=9,104				
Panel C: Black and Hispanic				
ELA	–0.295	0.510*	0.0166	–0.233**
	(0.320)	(0.292)	(0.347)	(0.107)
Math	0.547*	0.167	–0.423*	–0.292
N=40,731	(0.308)	(0.274)	(0.225)	(0.330)

Notes: This table reports IV estimates of the effect of the Asian-student share on the share of non-Asian students in four proficiency groups. Each cell reports a coefficient from a separate regression. All specifications include school, year, cohort, grade, grade-by-year, and race (except panel B) fixed effects. Regressions are weighted by the number of students completing the test. Standard errors (in parentheses) are clustered at the census-tract level. *** p<0.01, ** p<0.05, * p<0.10.

Online Appendix A
[Not for print publication unless otherwise requested]

Figure A1: Geocoded Schools (orange circle) and 1990 Census Tract Boundaries (blue line) in Manhattan, New York City

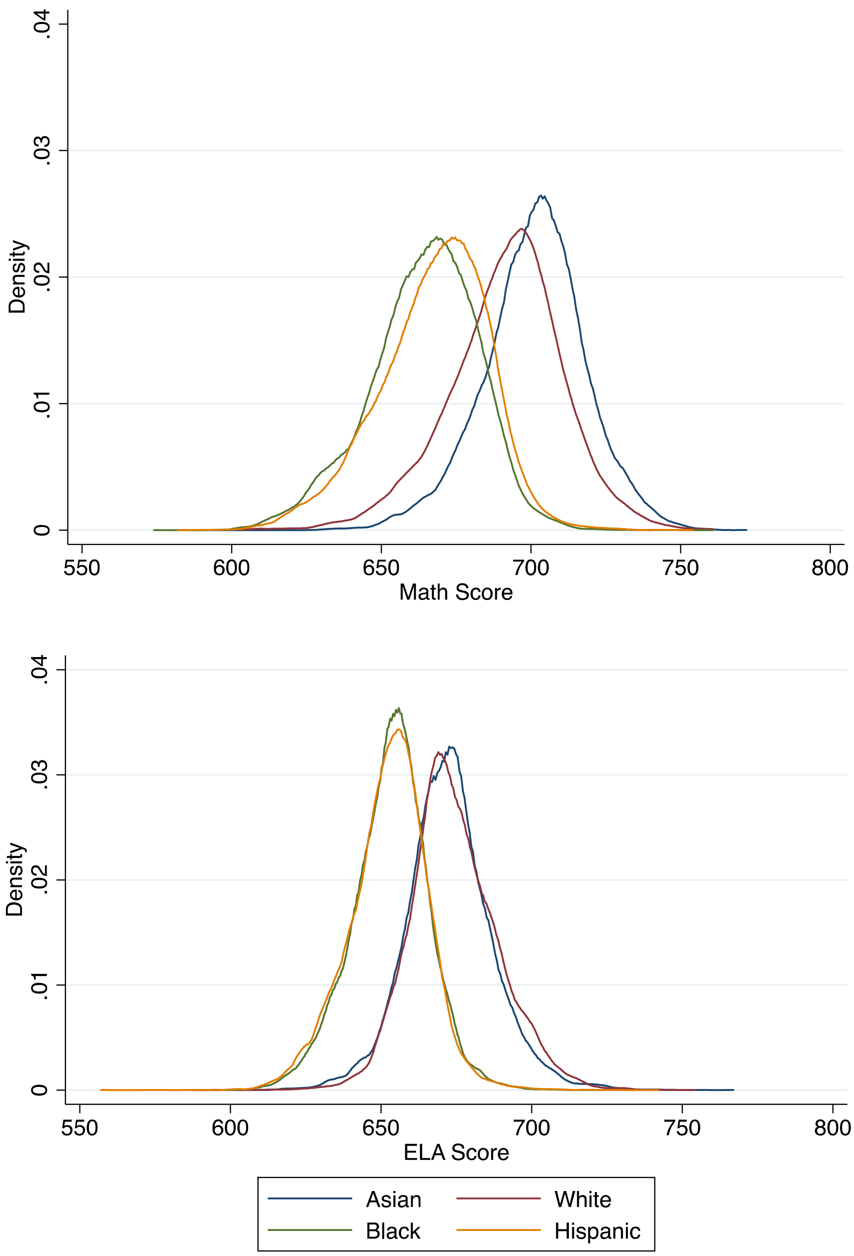

Figure A2: Test Score Distributions by Race/Ethnicity

Notes: This figure displays test score distributions for math and ELA by race/ethnicity. It is constructed from school-grade-year-race/ethnicity mean scores and by weighting by the number of students in each cell. The figures for student-year test score observations on which the data are based are 405,518, 400,527, 830,823, and 1,123,957 for Asian, white, black, and Hispanic students respectively. Data source: New York City Department of Education mean test score files by school, grade, year, and ethnicity/race for the years 2006 through 2012.

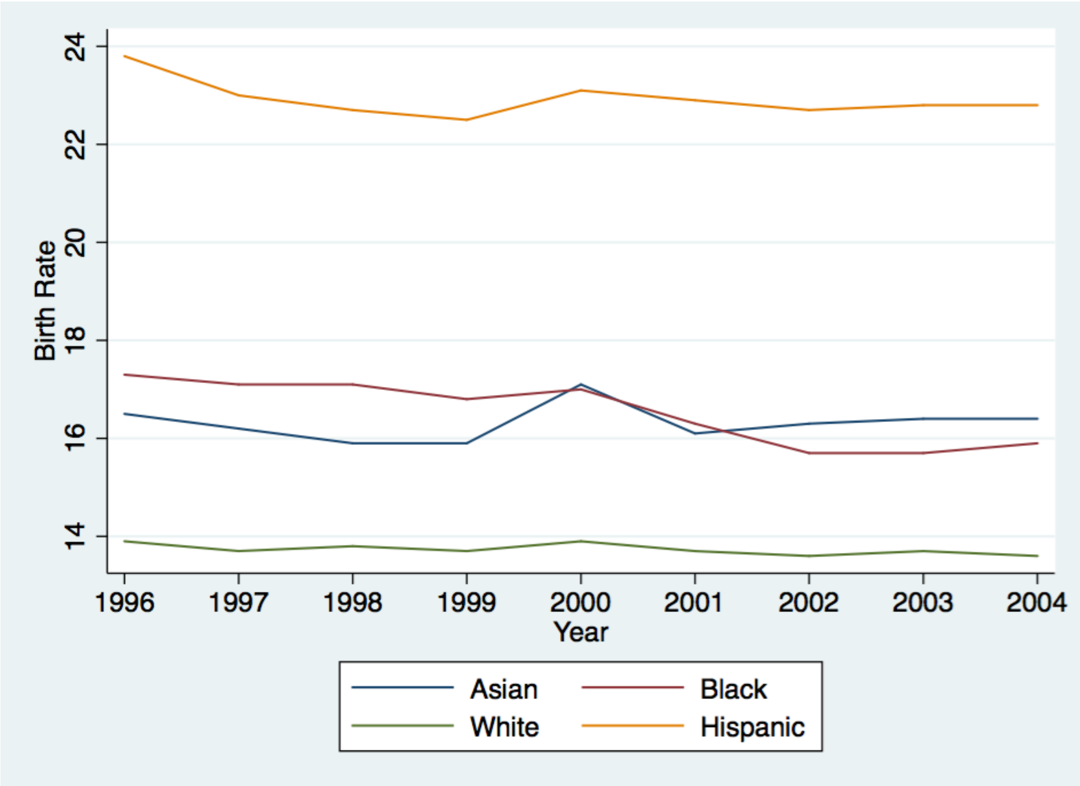

Figure A3: Birth Rates per 1,000 Individual in the U.S. by Race/Ethnicity

Notes: This figure displays annual birth rates per 1,000 individuals by race/ethnicity and year. Data source: U.S. National Center for Health Statistics.

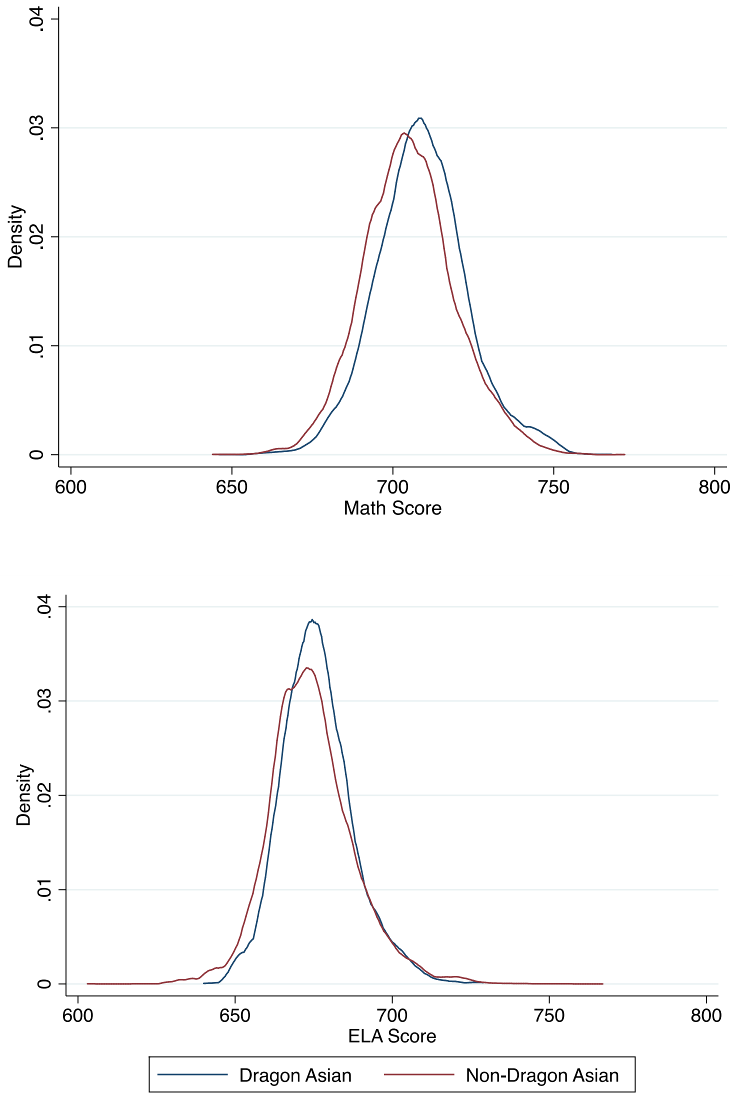

Figure A4: Asian Test Score Distributions

Notes: This figure displays Asian test score distributions for math and ELA separately for Dragon and non-Dragon cohorts. The non-Dragon cohorts used are the cohorts born three years before (1997–1999) and three years after (2001–2013) the Dragon cohort. Test score distributions are constructed from Asian school-grade-year score means and are weighted by the number of students completing the test. Data source: New York City Department of Education mean test score files by school, grade, year, and ethnicity/race for the years 2006 through 2012.

Table A1
Joint Distribution of Minimum and Maximum Grade

Minimum Grade	Maximum Grade: K	1	2	3	4	5	6	7	8	9	10	11	12	Total
PK	2	5	13	25	35	2,080	322	52	428	12	6	1	1	2,982
K	1	3	6	52	80	1,678	233	26	399	5	40	1	17	2,541
1	0	0	0	0	0	4	3	0	0	0	0	0	1	8
2	0	0	0	0	6	17	0	0	0	0	1	0	0	24
3	0	0	0	0	1	24	11	0	2	1	0	0	1	40
4	0	0	0	0	1	16	0	0	8	1	1	1	8	36
5	0	0	0	0	0	2	1	0	132	7	13	3	0	158
6	0	0	0	0	0	0	55	92	1,415	106	116	41	280	2,105
7	0	0	0	0	0	0	0	1	24	18	7	3	56	109
8	0	0	0	0	0	0	0	0	2	0	0	0	1	3
9	0	0	0	0	0	0	0	0	0	2	2	2	10	16
Total	3	8	19	77	123	3,821	625	171	2,410	152	186	52	375	8,022

Notes: This table reports the joint distribution of the minimum and maximum grades in a school (data from 2005 through 2012). “PK” indicates pre-kindergarten. “K” indicates kindergarten.

Table A2
Fraction of Students in Proficiency Groups by Ethnicity/Race (%)

	(1)	(2)	(3)	(4)
Proficiency Group:	1st Group	2nd Group	3rd Group	4th Group
		Panel A. Math		
Asian	2.5	9.8	39.2	48.3
White	3.6	14.5	45.4	36.3
Black	10.1	32.1	45	12.6
Hispanic	9	28.3	46.5	16
		Panel B. ELA		
Asian	4.4	21.9	61.9	11.7
White	4.6	23.5	60.5	11.2
Black	10.2	43.2	43.8	2.7
Hispanic	10.8	40.7	45.2	3.2

Notes: This table shows the fraction of students in each proficiency group by ethnicity/race. The proficiency categories are well below proficient (1st), below proficient (2nd), proficient (3rd), and above proficient (4th).

Table A3

Summary Statistic for School-Year-Level Variables

	Full Sample			1990 Chinese Population Share: Below Median			1990 Chinese Population Share: Above Median		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
% Free Lunch	69.1	23.8	7,153	74.1	22.1	4,096	62.3	24.4	3,057
% Reduced-Price Lunch	8.2	5.3	7,171	6.9	4.7	4,106	10	5.4	3,065
% Attendance	92	6.5	7,159	90.9	7.8	4,095	93.5	3.8	3,064
% Suspended	3.3	4.8	7,160	3.9	5.2	4,096	2.5	4.1	3,064
Number of Teachers	47.4	21.7	7,157	44.4	20	4,098	51.5	23.1	3,059
% Teachers without Valid Certification	2.6	4.3	7,157	3.3	4.9	4,098	1.8	3.1	3,059
% Teachers out of Certification	8.3	8.5	7,157	9.5	9	4,098	6.7	7.6	3,059
% Teachers with Less Than 3 Years' Experience	13.1	12.5	7,157	14.6	13.7	4,098	11.2	10.3	3,059
Pupils per Teacher	8.2	3.8	7,157	8.3	3.7	4,098	8.2	3.9	3,059
Total Number of Classes Taught	100.5	74.8	7,157	90.7	62.1	4,098	113.6	87.3	3,059
% Classes with Teacher without Appropriate Certification	10.1	9.9	7,156	11.3	10.5	4,098	8.3	8.7	3,058
Average Class Size	24.4	3.9	5,614	24.3	3.9	3,171	24.6	3.8	2,443
Index of Non-Asian Fractionalization (2006)	0.37	0.18	964	0.34	0.17	548	0.41	0.17	416

Notes: Descriptive statistics at the school-year level (data from 2006 through 2012). Index of non-Asian fractionalization is computed in 2006, the first year in which we observe attendance by race/ethnicity. Descriptive statistics are all reported for the full sample and for samples below and above the median of the local 1990 Chinese population share.

Table A4

Chinese and Asian Population Shares in 1990

			1990 Chinese Population Share			
	Full Sample		Below Median		Above Median	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Chinese % (500 m)	2.7	(6.3)	0.5	(0.4)	5.9	(8.7)
Chinese % (1,000 m)	2.9	(5.6)	0.6	(0.5)	6.0	(7.5)
Chinese % (2,000 m)	2.8	(4.3)	0.8	(0.9)	5.7	(5.3)
Chinese % (3,000 m)	2.7	(3.2)	1.1	(1.3)	5.0	(3.6)
All Asian % (500 m)	5.3	(7.8)	1.8	(2.0)	10.2	(9.9)
All Asian % (1,000 m)	5.5	(7.1)	2.0	(1.9)	10.3	(8.7)
All Asian % (2,000 m)	5.6	(5.9)	2.4	(1.9)	9.9	(6.8)
All Asian % (3,000 m)	5.5	(4.9)	2.9	(2.1)	9.1	(5.4)
Asian Excl. Asian Indian % (500 m)	4.3	(7.2)	1.1	(1.2)	8.6	(9.4)
Asian Excl. Asian Indian % (1,000 m)	4.4	(6.5)	1.3	(1.2)	8.7	(8.2)
Asian Excl. Asian Indian % (2,000 m)	4.5	(5.3)	1.6	(1.4)	8.3	(6.3)
Asian Excl. Asian Indian % (3,000 m)	4.4	(4.3)	2.0	(1.8)	7.6	(4.8)
Number of Schools	1,081		625		456	

Notes: Descriptive statistics for local 1990 Chinese and Asian population shares. The variables are constructed by using a GIS procedure to find census tracts within 500 to 3,000 meters of a school. When several nearby census tracts are identified, the population weighted average of ethnic population shares are used. Measures for groups including Chinese, Asians excluding Asian Indians, and all Asians are reported for the full sample and for samples below and above the median of the local 1990 Chinese population share. The unit of observation is a school.

Table A5

Impact of the Instrument on Asian Test Scores and Grade-Level Average Class Size, All Grades

	(1)	(2)	(3)
Dependent Variable:	Asian Math Score	Asian ELA Score	Grade-Level Average Class Size
% Chinese 1990 × Dragon	0.0099	0.0110	0.0062
	(0.0238)	(0.0268)	(0.0086)
Observations	9,478	9,478	14,455
Number of Schools	542	542	1,065

Notes: This table reports reduced-form estimates of the effect of the instrument on Asian test scores and grade-level average class size using data for all grades. The unit of observation is a school-grade-year. All regressions include school, cohort, and grade-by-year fixed effects. The instrument is the interaction between the Dragon dummy and local 1990 Chinese population share. Standard errors (in parentheses) are clustered at the census-tract level. *** p<0.01, ** p<0.05, * p<0.10.

Table A6
Robustness Checks: Trimming the Sample by Initial Enrollment and Local 1990 Chinese Population Share

	(1)	(2)	(3)	(4)
	Initial Enrollment Excluding Top Percentiles		1990 Chinese Share Excluding Top Percentiles	
ELA Non-Asian	–0.818***	–0.875***	–0.706**	–0.728**
	(0.305)	(0.309)	(0.345)	(0.334)
Math Non-Asian	–0.647**	–0.684**	–0.580**	–0.557**
	(0.300)	(0.309)	(0.282)	(0.264)
Observations	49,445	49,002	49,487	49,115
Schools	1,073	1,065	1,070	1,060

Notes: This table reports IV estimates of the effect of the share of Asian students on non-Asian ELA and math test scores. All specifications control for school, cohort, year, grade, grade-by-year, and race fixed effects. Columns 1 and 2 exclude observations above the 99th and 98th percentiles of initial enrollment measured in the first year the school is observed in the data respectively. Columns 3 and 4 exclude observations above the 99^{th} and 98^{th} percentiles of the local 1990 Chinese population share respectively. Regressions are weighted by the number of students completing the test. Standard errors (in parentheses) are clustered at the census-tract level. *** p<0.01, ** p<0.05, * p<0.10.

Table A7

Estimates Based on Alternative Definitions of the Instrument

	(1)	(2)	(3)	(4)	(5)	(6)
	ELA, Non-Asian Students			Math, Non-Asian Students		
	First Stage	Reduced Form	IV	First Stage	Reduced Form	IV
% Chinese 1990 (500 m) × Dragon	0.151**	–0.116***	–0.765**	0.153**	–0.0979**	–0.639**
	(0.0600)	(0.0447)	(0.314)	(0.0603)	(0.0438)	(0.311)
% Chinese 1990 (1,000 m) × Dragon	0.146**	–0.117**	–0.800**	0.148**	–0.0651	–0.440
	(0.0577)	(0.0454)	(0.370)	(0.0582)	(0.0420)	(0.280)
% Chinese 1990 (2,000 m) × Dragon	0.195***	–0.175***	–0.900***	0.197***	–0.0773*	–0.393
	(0.0501)	(0.0491)	(0.343)	(0.0505)	(0.0453)	(0.239)
% Chinese 1990 (3,000 m) × Dragon	0.274***	–0.284***	–1.035***	0.277***	–0.0992*	–0.358
	(0.0510)	(0.0637)	(0.292)	(0.0514)	(0.0584)	(0.219)
% Asian Excl. Indians 1990 (500 m) × Dragon	0.154***	–0.114***	–0.741**	0.155***	–0.0851***	–0.548**
	(0.0409)	(0.0384)	(0.303)	(0.0410)	(0.0318)	(0.236)
% Asian Excl. Indians 1990 (1,000 m) × Dragon	0.150***	–0.113***	–0.750**	0.151***	–0.0650**	–0.430*
	(0.0396)	(0.0377)	(0.316)	(0.0398)	(0.0318)	(0.220)
% Asian Excl. Indians 1990 (2,000 m) × Dragon	0.167***	–0.148***	–0.887***	0.168***	–0.0690**	–0.411*
	(0.0323)	(0.0388)	(0.297)	(0.0324)	(0.0335)	(0.212)
% Asian Excl. Indians 1990 (3,000 m) × Dragon	0.197***	–0.203***	–1.031***	0.198***	–0.0814**	–0.412**
	(0.0361)	(0.0489)	(0.289)	(0.0363)	(0.0379)	(0.205)
% Asian 1990 (500 m) × Dragon	0.145***	–0.0979***	–0.673**	0.146***	–0.0926***	–0.633***
	(0.0319)	(0.0328)	(0.270)	(0.0320)	(0.0289)	(0.239)
% Asian 1990 (1,000 m) × Dragon	0.146***	–0.100***	–0.686**	0.147***	–0.0795***	–0.542**
	(0.0315)	(0.0338)	(0.280)	(0.0316)	(0.0299)	(0.222)
% Asian 1990 (2,000 m) × Dragon	0.153***	–0.128***	–0.836***	0.154***	–0.0810***	–0.525**
	(0.0275)	(0.0356)	(0.275)	(0.0277)	(0.0310)	(0.219)
% Asian 1990 (,3000 m) × Dragon	0.170***	–0.170***	–0.999***	0.171***	–0.0936***	–0.547**
	(0.0319)	(0.0439)	(0.282)	(0.0321)	(0.0337)	(0.217)

Notes: This table reports IV estimates of the effect of the instrument on non-Asian ELA and math test scores based on alternative definitions of the instrument, and the corresponding first-stage and reduced-form coefficients on the instrument. All specifications include school, cohort, year, grade, grade-by-year, and race fixed effects. The number of observations is 49,972 in each regression. Results are reported for the baseline specification (first row) and for alternative school neighborhoods (census tracts within 500, 1000, 2000, and 3000 meters of a school) and Asian groups (Chinese, Asians excluding Indians, and all Asians). Regressions are weighted by the number of students completing the test. Standard errors (in parentheses) are clustered at the census-tract level. *** p<0.01, ** p<0.05, * p<0.10.

Table A8
Effects by Race
(First-Stage, Reduced-Form, and IV Estimates)

	(1) White	(2) Hispanic	(3) Black	(4) Hispanic-Black
Panel A: First Stage				
ELA	0.221***	0.148**	0.0763	0.135**
	(0.0636)	(0.0696)	(0.0527)	(0.0639)
Math	0.219***	0.150**	0.0781	0.136**
	(0.0632)	(0.0702)	(0.0531)	(0.0643)
Panel B: Reduced Form				
ELA	0.0834	–0.0877**	–0.107	–0.0769**
	(0.0944)	(0.0341)	(0.0733)	(0.0339)
Math	–0.0989	–0.0842*	–0.0691	–0.0748*
	(0.0875)	(0.0439)	(0.0754)	(0.0440)
Panel C: IV				
ELA	0.377	–0.591*	–1.409	–0.571*
	(0.454)	(0.322)	(0.992)	(0.295)
Math	–0.451	–0.560*	–0.884	–0.549
	(0.399)	(0.329)	(1.126)	(0.350)
Observations	9,104	21,892	18,795	40,730
Number of Schools	501	1,051	956	1,077

Notes: First-stage, reduced-form, and IV estimates corresponding to specifications in Table 5, panel A. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table A9
Effects by Grade among Non-Asian Students
(First-Stage, Reduced-Form, and IV Estimates)

	(1) 3rd	(2) 4th	(3) 5th	(4) 6th
Panel A: First Stage				
ELA	0.112**	0.148**	0.163**	0.124
	(0.0561)	(0.0601)	(0.0650)	(0.0775)
Math	0.115**	0.150**	0.166**	0.126
	(0.0566)	(0.0610)	(0.0658)	(0.0777)
Panel B: Reduced Form				
ELA	–0.0826*	–0.0573	–0.187**	–0.0741
	(0.0422)	(0.0696)	(0.0778)	(0.0556)
Math	–0.143**	–0.0595	–0.127*	–0.0610
	(0.0642)	(0.0637)	(0.0738)	(0.0684)
Panel C: IV				
ELA	–0.741*	–0.386	–1.149**	–0.596
	(0.406)	(0.460)	(0.540)	(0.572)
Math	–1.240**	–0.396	–0.764*	–0.484
	(0.532)	(0.379)	(0.434)	(0.754)
Observations	9,982	9,937	9,930	7,216
Number of Schools	722	715	731	598

Notes: First-stage, reduced-form, and IV estimates corresponding to specifications in Table 5, panel B. *** p<0.01, ** p<0.05, * p<0.10.

Table A10

Heterogeneity by School Characteristics, Additional Results

Outcome:	(1) Non-Asian ELA Score	(2) Non-Asian Math Score
Panel A. 2006 Average Test Score × % Chinese 1990 × Dragon	0.00166	0.00131
N=40,916	(0.00176)	(0.00210)
Panel B. % Teachers No Valid Cert. × % Chinese 1990 × Dragon	0.0142	0.0157
N=42,860	(0.0150)	(0.0130)
Panel C. % Teachers Out of Cert. × % Chinese 1990 × Dragon	0.00504	0.00923
N=42,860	(0.00534)	(0.00642)
Panel D. % Teachers Experience < 3 Yrs. × % Chinese 1990 × Dragon	0.00300	0.00161
N=42,860	(0.00306)	(0.00190)
Panel E. % Students with Free Lunch × % Chinese 1990 × Dragon	–0.00161	0.00112
N=45,812	(0.00289)	(0.00180)
Panel F. % Students with Reduced-Price Lunch × % Chinese 1990 × Dragon	–0.00293	0.00175
N=45,812	(0.00918)	(0.00915)

Notes: This table reports reduced-form estimates for the coefficient on the instrument interacted with a school characteristic. School characteristics are measured in the academic year of 2004/2005, whereas the average test score is measured in 2006. In all cases, this is the first year in which we observe these variables in our data. Each table cell reports a coefficient from a separate regression. For example, the estimate in the first row of column 1 is the coefficient on the interaction between the 1990 Chinese population share, Dragon dummy, and average test score for ELA. In panel A, we exclude the year of 2006 to avoid a mechanical correlation between the dependent variable and the 2006 school-level average test score. All specifications control for the main effects and the interaction between the Dragon dummy and school characteristic, the interaction between the Dragon dummy and 1990 Chinese population share, and the interaction between the school characteristic and 1990 Chinese population share, and they include school, cohort, grade-by-year, and race fixed effects. Regressions are weighted by the number of students attending the test. Standard errors (in parentheses) are clustered at the census-tract level. *** p<0.01, ** p<0.05, * p<0.10.

Table A11

Controlling for School- and Grade-Level Variables, IV Estimates

	(1)	(2)	(3)	(4)
	ELA Test Score		Math Test Score	
	Baseline	+ Controls	Baseline	+ Controls
Panel A. Without Grade-Level Average Class Size				
Asian Share	–0.781***	–0.632**	–0.622**	–0.547**
	(0.297)	(0.262)	(0.265)	(0.245)
Observations	39,735	39,735	39,735	39,735
Number of Schools	912	912	912	912
Panel B. With Grade-Level Average Class Size (Excludes the Academic Year 2006/2007)				
Asian Share	-0.594**	-0.495*	-0.569**	-0.503**
	(0.268)	(0.257)	(0.240)	(0.240)
Observations	32,971	32,971	32,971	32,971
Number of Schools	894	894	894	894

Notes: This table reports IV estimates of the effect of the Asian-student share on non-Asian ELA and math test scores. All specifications control for school, cohort, year, grade, grade-by-year, and race fixed effects. In Panel A, columns 1 and 3 display estimates for the baseline specification that correspond to IV estimates in Table 3 but are based on a restricted sample of schools for which control variables included in columns 2 and 4 are available. Columns 2 and 4 add the students per teacher ratio; share of white students; share of Hispanic students; attendance rate; share of suspended students; share of students eligible for free and reduced-price lunch; number of teachers; share of teachers without valid certification, out of certification, and with less than three years of experience; number of classes; share of classes taught by teachers without appropriate certification; and average number of students in self-contained classes in first through sixth grade. Panel B provides results for corresponding specifications with grade-level average class size added to the set of control variables. This further restricts the sample size because the variable is unavailable for the academic year 2006/2007. Regressions are weighted by the number of students completing the test. Standard errors (in parentheses) are clustered at the census-tract level. *** p<0.01, ** p<0.05, * p<0.10.

Table A12
Effects on Proficiency, Non-Asian Students
(First-Stage, Reduced-Form, and IV Estimates)

Proficiency Level:	(1) 1st	(2) 2nd	(3) 3rd	(4) 4th
Panel A: First Stage				
ELA	0.151**	0.151**	0.151**	0.151**
	(0.0600)	(0.0600)	(0.0600)	(0.0600)
Math	0.153**	0.153**	0.153**	0.153**
	(0.0603)	(0.0603)	(0.0603)	(0.0603)
Panel B: Reduced Form				
ELA	–0.0425	0.0687*	0.0239	–0.0502**
	(0.0440)	(0.0372)	(0.0481)	(0.0205)
Math	0.0822**	0.0169	–0.0562*	–0.0431
	(0.0319)	(0.0342)	(0.0311)	(0.0404)
Panel C: IV				
ELA	–0.281	0.454*	0.158	–0.332**
	(0.258)	(0.242)	(0.303)	(0.136)
Math	0.537**	0.111	–0.367*	–0.282
	(0.265)	(0.229)	(0.214)	(0.295)
Observations	49,872	49,872	49,872	49,872
Number of Schools	1,080	1,080	1,080	1,080

Notes: First-stage, reduced-form, and IV estimates corresponding to specifications in Table 9, panel A. *** p<0.01, ** p<0.05, * p<0.10.

A.1 Math and ELA tests scores and Grade-Level Average Class Size

Math and ELA tests are standardized exams conducted in the spring semester in third through eighth grade. The math tests cover the following topics: (1) number sense and operations, (2) algebra, (3) geometry, (4) measurement, and (5) statistics and probability. Tests in the earlier grades emphasize basic content, such as number sense and operations, whereas tests in the later grades focus on more advanced topics, such as algebra and geometry. The ELA tests are designed to assess students in three learning standards: (1) information and understanding, (2) literary response and expression, and (3) critical analysis and evaluation. The ELA tests include multiple-choice and short-response sections and reading and listening exercises, as well as brief editing tasks.

The number of correct answers is converted into a "scale score." The aim of the scaling is to improve the comparability of scores across grades and years. For instance, the DOE considers the difficulty of the question, its capacity to differentiate between high- and low-performing students, and the likelihood of getting a correct answer by guessing. Scale scores are divided into four performance levels: NYS Level 1 (well below proficient), NYS Level 2 (below proficient), NYS Level 3 (proficient), and NYS Level 4 (above proficient). Thresholds for these proficiency levels are determined annually for each grade at the state level by a panel of experts.

Data on average class size at the school-grade-year level were downloaded from the DOE website.[40] Annual files contain information on average class size from academic year 2006/2007. Data on year 2005/2006 are unavailable. For academic years 2006/2007 and 2007/2008, only one data file exists. From 2008/2009 onwards two files for each year are provided: preliminary and updated version. We use the updated version, which contains officially audited information. We then match class size information at the school-grade-year level with the main data used for the empirical analysis. The match can be implemented via school identifiers starting from 2007/2008. No consistent school identifier is provided in the 2006/2007 data. These information are hence merged via a school name matching procedure (98% match rate) and a final manual search, which allowed us to match all remaining observations.

A.2 New York State School Report Card and Grade-Level Class Size Data

[40] http://schools.nyc.gov/AboutUs/schools/data/classsize/Class+Size+Archive.htm

School report cards for schools in the state of New York are publicly available at the following website: https://data.nysed.gov/downloads.php. Annual files contain various sub-files. We used the following files: "BEDS Day Enrollment," "Average Class Size," "Demographic Factors," "Attendance and Suspensions," and "Staff." Report cards are consistently available from the academic year of 2004/2005 onward. Each annual file contains information for the current academic year as well as for two or three previous academic years. If values for year t were missing in the year t file, we used the value for the year t in the year $t+1$ or $t+2$ files, if available. For example, a variable for the academic year of 2006/2007 might not be recorded in the 2007 file but might be recorded in the 2008 or 2009 files. In the paper, we use the following variables:

Name of the Variable	Definition
Common_branch	Number of students in self-contained classes in first through sixth grade divided by the number of such classes (average class size in self-contained classes)
Per_free_lunch	Fraction of enrolled students eligible for free lunch
Per_reduced_lunch	Fraction of enrolled students eligible for reduced-price lunch
Num_teach	Total number of teachers
Per_no_valid_cert	Fraction of teachers with no valid teaching certificate
Per_teach_out_of_cert	Fraction of teachers out of certification
Per_fewer_three_years_exp	Fraction of teachers with fewer than three years of teaching experience
Num_class	Total number of classes taught
Per_no_appropriate_cert	Fraction of classes taught by teachers without appropriate certification
Attendance_rate	Annual attendance rate
Per_suspension	Fraction of suspended students

A.3 Geocoding of Schools

We used school addresses drawn from the New York State School Report Cards to geocode schools. The first step of our geocoding procedure used the U.S. Census Bureau address batch geocoder.[41] We were able to obtain coordinates for 976 schools (90.3%). The resulting address matches and co-ordinates were manually checked using Google's map service and

[41] https://www.census.gov/geo/maps-data/data/geocoder.html

schools' websites.[42] We conducted a Google search for the remaining 108 schools for which the batch geocoder did not provide results: 105 of them were found in the address reported in the report card. Coordinates for these schools were recorded using their locations in the Google map service, whereas the remaining three schools were excluded from the analysis.

A.4 Population Shares by Ethnicity/Race in a School Neighborhood

We constructed variables for the historical ethnic/racial structure of population in a school's neighborhood using 1990 census data on population by ethnic/racial group by census tract provided by the Minnesota Population Center.[43] Census tracts were linked to schools in the ArcMap program by overlaying school coordinate points on a shapefile of 1990 census-tract boundaries provided by the Minnesota Population Center (Adams et al., 2004) and finding census tracts within a given distance of the school. Figure A1 displays an example of geocoded schools and census-tract boundaries in Manhattan, a borough of New York City. To identify school neighborhoods, we ran an ArcMap procedure to list census tracts that were within the distance of 500, 1,000, 2,000, and 3,000 meters of the school. We calculated population shares in a school's neighborhood as the fraction of an ethnic population (e.g., Chinese or Asian) to the total population of the school neighborhood.

[42] We detected 16 errors in the address batch-geocoding results. Many of these were due to the relatively complicated NYC street address system. For instance, some errors were associated with addresses including the word "East" and the abbreviation "E" for it. Another common reason for an error was street numbering based on intersecting major streets.
[43] We used the file nhgis0001_ds120_1990_tract.