Journal of Urban Economics 77 (2013) 27–40

Contents lists available at SciVerse ScienceDirect

Journal of Urban Economics

www.elsevier.com/locate/jue

# Do small schools improve performance in large, urban districts? Causal evidence from New York City

Amy Ellen Schwartz [a,c,*], Leanna Stiefel [a,c], Matthew Wiswall [b,c]

[a] Steinhardt and Wagner Schools, New York University, 295 Lafayette Street, Third Floor, New York, NY 10012, USA
[b] Department of Economics, W.P. Carey School of Business, Arizona State University, PO Box 879801, Tempe, AZ 85287-9801, USA
[c] Institute for Education and Social Policy, New York University, USA

ARTICLE INFO

*Article history:*
Received 27 July 2012
Revised 6 February 2013
Available online 10 April 2013

*JEL classification:*
I2
H4

*Keywords:*
Education reform
Small schools
Heterogeneous treatment
Instrumental variables

ABSTRACT

We evaluate the effectiveness of small high school reform in the country's largest school district, New York City. Using a rich administrative dataset for multiple cohorts of students and distance between student residence and school to instrument for endogenous school selection, we find substantial heterogeneity in school effects: newly created small schools have positive effects on graduation and some other education outcomes while older small schools do not. Importantly, we show that ignoring this source of treatment effect heterogeneity by assuming a common small school effect yields a misleading zero effect of small school attendance.



## 1. Introduction

While the academic achievement of US elementary school students has improved over the last decade, US high school students continue to graduate at unacceptably low rates and measures of achievement show only a slight upward trend since 2005 (National Center for Education Statistics, 2010). Moreover, the achievement and graduation gaps between white and black high school students and between white and Hispanic high school students have not changed. For example, by some calculations slightly over 80% of white students graduate within 4 years, but only 60% of black and 62% of Hispanic students do so and the gap in college readiness is similarly stark. In addition, with the erosion of the labor market for low-skilled workers over the past several decades, the gap in earnings between high school graduates and non-graduates has increased. Within this context many school systems with large proportions of poor students, in particular large urban school systems, face tremendous challenges; a majority of their students are at risk of not succeeding in high school and thus have more limited access to post-secondary education and have lower labor market earnings than many of their counterparts in suburban districts. While several reforms target high school students, the *small school reform* stands out because of its adoption in many major cities and its substantial public and philanthropic funding base. Placing students in small schools is advocated as a way to provide students with the support they need to improve their performance.

There have been at least two waves on small high school reform in US cities as well as an early and more recent literature on their effectiveness. The early wave of small school reforms in cities such as New York City (NYC), Chicago, Philadelphia and Oakland occurred in the early 1990s; and many of the intentionally-small high schools created then still exist. The early literature that stimulated and accompanied these reforms was conceptual (establishing theoretical reasons why small schools would help disadvantaged youth) and, when empirical, correlational in nature. A later wave of small school reform occurred after 2000 in NYC, Chicago, Los Angeles, San Diego, Philadelphia, and Boston among others, often with some funding from large foundations such as the Gates Foundation, the Carnegie Corporation, and the Open Society Institute as well as the U.S. Department of Education (U.S. Department of Education, 2006). The literature on this wave includes some studies using regression analysis (Stiefel (2009) for example), and, in one case, a lottery design (Bloom et al. 2010).

* Corresponding author at: Steinhardt and Wagner Schools, New York University, 295 Lafayette Street, Third Floor, New York, NY 10012, USA. Fax: +1 212 995 4564.
*E-mail addresses:* amy.schwartz@nyu.edu (A.E. Schwartz), leanna.stiefel@nyu.edu (L. Stiefel), matt.wiswall@gmail.com (M. Wiswall).

0094-1190/$ - see front matter 
http://dx.doi.org/10.1016/j.jue.2013.03.008

In this study, we use administrative data covering all NYC public high school students to evaluate the effectiveness of two generations of small schools in NYC. The long recognized challenge in educational evaluations is the possible selection of students into the education intervention, which can bias simple comparisons of outcomes for those who are treated by the intervention and those who are not. In our application, selection bias arises if students who attend small and large schools differ on dimensions, such as motivation, ability, and parental support, which have an independent effect on the outcomes of interest. We address selection bias in two ways. First, we use a rich set of student characteristics, such as gender, race, language skills, and prior middle school test scores, to control for many of the observable differences between students attending small and large high schools. But, as in a wide variety of evaluation contexts, the observed student characteristics in our data are unlikely to fully eliminate unobserved or unmeasured differences in student characteristics that affect student outcomes.

Recognizing this potentially important selection on unobservables, we next turn to quasi-experimental methods using credible instrumental variables that exogenously influence student decisions to attend small schools but do not influence student outcomes. Since high schools of various sizes are not evenly distributed across the city, and students who live in the immediate vicinity of a small high school (especially relative to a large school) are more likely to attend a small school, we use as instruments the distance between the nearest small school or large school and the student's home.

Motivating our use of distance as an instrumental variable is a small but growing literature on the determinants of school choice.[1] A consistent result in the literature is that location (and specifically distance) of a school relative to a student's home residence is an important variable for students and parents in their choice of school. Schneider and Buckley (2002) report that in parent internet search behavior, location is the second most sought after piece of information after school demographics. Burgess and Briggs (2010), in a study of parental preferences for schools in England, conclude that parents make tradeoffs among academic attainment, school socio-economic composition, and travel distance. Hastings et al. (2006) find that in North Carolina proximity is highly valued by all, although families with strong preferences for academics are generally willing to tolerate longer distances. Saporito and Lareau (1999) conclude that both whites and blacks tend to choose schools close to their homes but whites are often willing to travel further to attend schools with higher proportions of white students. Motivated by this prior literature, we form instruments from the distance between the nearest small or large school and the student's home. A similar instrumental variables framework has been used in an educational evaluation of Chicago schools (Cullen et al., 2005), an evaluation of small schools (Barrow et al., 2010) and charter schools (Booker et al., 2011) in Chicago, and an examination of the effect of college attendance on earnings (Card, 1995) and on health behaviors (Currie and Morretti 2003). As this prior research has demonstrated in a variety of contexts, the likelihood of attending a school decreases as the distance to the school increases, perhaps because of higher costs such as those involving transportation.

We confirm these results with our NYC data and show that distance strongly predicts actual small school attendance, even after conditioning on student characteristics. We also present several additional analyses that support the instrument exogeneity. We use these distance based IVs to instrument for small school attendance and obtain IV estimates of the effects of attending small schools. Suggesting the importance of student sorting into schools based on unobserved student characteristics such as motivation, we find a positive effect of small school attendance with the OLS estimator but a small and imprecise estimate using the IV estimator.

An important contribution of this paper is to distinguish between the old and new generation of small schools. Rather than assume a common small school effect, we instead divide the small schools into those newly developed since 2002 and those which existed prior to the latest wave of reforms. These new small schools are different in a number of ways from the old small schools, differences that we further explore to assess whether they are related to effectiveness and whether they can or will be sustained.

Using models where we distinguish between new and old small schools, we find important differences in the effects of the schools in both our OLS and IV estimates. In our IV estimates, using instruments for distance to the new and old small schools, attending an old small school is estimated to have a negative effect on the probability of graduating relative to large schools, while attending a new small school is estimated to increase graduation rates by 17% relative to attending a large school. This estimate is statistically significant from zero at the five percent level and the magnitude of the estimate is robust to changes in sample selection, variable definitions, and various alternative instrumental variable estimators. When we turn to other high school outcomes, we find more mixed results. Our IV estimates indicate that attending a new small school increases the probability of taking the Regents English and mathematics examinations by 14% and 16% respectively. We estimate, however, a negative effect of new small school attendance on English scores and no effect on mathematics scores. We cannot rule out, however, that the non-positive effects on test scores is the result of the marginal test takers induced to take these exams having lower ability. In addition, we estimate that old small schools have considerably more negative effects on test scores than the new small schools.

Our estimates reveal a clear divide in the effects of the new versus old small schools and provide some context to understand the results of previous research. Studying a subset of the new small schools, which were over-subscribed and offered admission by lottery, Bloom et al. (2010) find a 6.8% point increase in graduation rates from attending these "small schools of choice". Our positive statistically significant effect of new small school attendance estimated using a different empirical strategy – the IV – is consistent with this finding, and given the standard errors, both estimates are within the other's 95% confidence interval. An advantage of our study is that we can use our identification strategy to estimate the effect of a wider variety of small school types. Thus, our estimates show that the positive effects estimated for the recent small schools would not necessarily extrapolate to *all* small schools. In particular, the older generation of small schools is estimated to have a negative effect on graduation, in contrast to the positive effect of the newer generation. This is a crucial finding for policy: school size matters but it is not sufficient for affecting outcomes on its own. It also provides a cautionary tale for policymaking in general. Much of the original enthusiasm by foundations and districts for small school reform was based on early OLS studies relating size to outcome; prior to 2002 there were no causal studies of the effects of small high schools. The results of our study show that, while OLS estimation yields positive effects for the old small

[1] Studies that rely on survey analysis indicate that parents generally choose based on academic achievement and the quality of teachers at the school (Armour and Peiser, 1998; Greene et al.1998) and do not exhibit much preference towards student demographics (Schneider et al., 1998). In contrast to survey studies, observed choice behavior reveals that parents do have strong preferences for schools with similar demographics (Schneider and Buckley, 2002; Hastings et al. (2006)). There is also literature on the impacts of choice, for example on socioeconomic and racial sorting and segregation (Ladd, 2003) and on decisions between private and public school (Hoyt and Lee, 1998).

schools, these effects disappear and even become negative with IV estimation – which suggests the OLS results were biased. Correlational evidence alone is not enough to support major policy changes. The new small school reforms, in the end, involved more than size reductions, but the rhetoric emphasized the effect of "being small", an effect that is not confirmed with econometric methods.

The paper is organized as follows. In Section 2, we review the literature on small high school reform and situate our contribution in this literature. In Section 3, we describe the NYC context for our evaluation. In Section 4, we describe our data and measures, in Section 5, we present our models and methods, and in Section 6 we present results. In Section 7, we discuss why we might find our results and in Section 8, we conclude with a discussion of the relevance of the results for policymakers.

## 2. Previous literature and our contribution

### 2.1. What is Small?

It is important to note that there is no consensus on a definition of "small" in the literature on school size and outcomes. The federal government, through its *Small Schools Initiative,* set a limit of 300 students (U.S. Department of Education, 2006) while the Gates-funded initiative in NYC considered 500 students the upper limit for small high schools (Gootman, 2005) and a recent study in Chicago established a 600 student cutoff (Barrow et al., 2010). In previous research in the mid-1990s on the costs of small high schools in NYC, as well as the then-current local policy, schools with 600 students or fewer were considered small (Stiefel et al., 2000). Lee and Smith (1997) found schools in the range of 600–900 to be most effective for minority students, and finally, a recent Gates funded study (Bloom et al., 2010) used 550 as the cutoff for a small school. To incorporate a policy-relevant figure in the range of the literature, we focus first on a 550-student cutoff and then perform a sensitivity analysis to ascertain the effect of alternative small sizes.

### 2.2. Early literature on small high schools

Despite the lack of consensus on what constitutes a small school enrollment, the early literature proposed many hypotheses about how small size could affect student outcomes. Fowler (1992) and Page et al. (2002) advanced the idea that small schools have more student participation in extracurricular activities and better student and teacher attitudes. Others hypothesized that small schools are particularly effective for disadvantaged students as a result of their superior social aspects, high perceived expectations for all students, teacher and administrator abilities to nurture students' (higher) needs, and better student behavior (Barker and Gump, 1964; Lindsay, 1982).

But the literature is not unified on the directions of effects, as some authors have claimed that increased numbers of academic offerings and a social climate that is more accepting of diversity are more likely in large schools (Haller et al., 1990; Watt, 2003). Moreover, high school reform in the mid 20th century featured the substitution of large high schools for small ones (Conant, 1959).

Empirical work before 2000, which was based on correlations and ignored issues of selection, suggested that achievement scores and attendance rates were higher and dropout rates were lower in small schools compared to large schools (Fowler and Walberg, 1991; Fowler, 1992; Lee and Smith, 1997). One of the most rigorous studies of student outcomes found that an optimal school size with respect to maximizing student achievement ranged between 600 and 900 students (Lee and Smith, 1997).

Another focus of the early research was heterogeneity in the effects of small schools for particular sub-groups defined by observed student characteristics. According to Fowler and Walberg (1991) and Fowler's (1992) literature reviews, schools with fewer than about 1500 deliver superior outcomes for minority and poor youth. Large schools may also have particularly negative effects on disadvantaged students, and small schools may better serve disadvantaged students not only vis-à-vis absolute achievement levels (Howley et al., 2000), but also with respect to lessening of achievement gaps (McMillen, 2004). Some empirical analyses, though, suggested that small schools provide benefits to all student types and that the distribution of gains across socioeconomic status and race is more equitable in smaller schools (Lee and Smith, 1995). None of these studies, however, addressed the potential selection bias or endogeneity of the choice to attend a small school based on unobservable student ability, motivation, parental involvement, etc.

### 2.3. More recent literature

More recently, scholars have turned their attention to the causality issue and have expanded their use of statistical methods and experimental designs to address it. Schneider et al. (2007) evaluated small school effects using data from the Educational Longitudinal Study of 2002 (ELS: 2002). They compared estimates from a random coefficients/hierarchical linear model (HLM) to those from a propensity score matching estimator using the available observable covariates. The authors found with both methods that attending a small high school has little effect on achievement, with the HLM estimates showing somewhat larger effects than the matching estimates for post-secondary expectations and number of colleges to which students applied. Both the HLM and matching frameworks have shortcomings in addressing selection, however. HLM, a type of control function approach, assumes a particular specification of the nested structure for outcomes. This approach, like parametric control function approaches, is not robust to model mis-specification. The matching framework, while potentially more flexible than some standard OLS regression frameworks, assumes that selection into small schools occurs only through the observable covariates available in the particular dataset used. Selection based on unobservables could bias these results.[2]

Barrow et al. (2010) used quasi-experimental variation in the distance between students' homes to high schools in an instrumental variables (IV) framework to evaluate the effect on performance of attending small high schools in Chicago. In their IV results, they found a positive effect of small school attendance on continuation through high school and graduation, but their study included only 22 small high schools and could not distinguish among school vintages.

Focusing on what Bloom et al. (2010) labeled "small schools of choice (SSC)", which enroll fewer than 550 students, the authors took advantage of lotteries for some seats in oversubscribed schools instituted by NYC's high school application processing system starting with the class entering ninth grade in 2005. Over the years 2005 through 2008, 105 such schools, serving students in grades 9–12 and founded after 2002, were part of the study's sample. Students not lotteried into a particular school might have been lotteried into another SSC. Therefore, the students not lotteried into any SSC and who attended another NYC high school, not necessarily a large one, served as the control group. Students in one

[2] See Heckman and Navarro (2004) for a discussion of the difference in identifying assumptions for matching versus control function approaches.

cohort (2005) were followed for four years and had statistically significantly higher graduation rates (6.8% higher) if they enrolled in SSC's than if they did not. Other cohorts, not in the schools for four years at the time of the study, had more credits toward graduation each year in SSCs than in other schools. The study provides some limited evidence on the effectiveness of a select group of small high schools – ones that were newly formed after 2002 and that were oversubscribed for some of their seats – that is, popular, new small schools. It does not provide evidence, however, on the broader issue of whether size itself is the critical feature of these schools. Since their sample includes neither older small schools nor undersubscribed small schools, it provides no insight into whether *small* schools are "better," rather than *new* schools or oversubscribed schools.[3] Finally, it does not provide a clear counterfactual since students may have attended other small schools if not lotteried into their choices.

In summary, our contributions involve the use of rich student- and school-level administrative data to study the effects of small high school attendance across multiple cohorts of students, attending several generations and types of small schools, the use of IV methods to obtain unbiased estimates, sensitivity to the definition of size, and the evaluation of multiple outcomes.

## 3. Context of small school reform in NYC

New York City is a particularly useful and relevant setting to study the effectiveness of small high schools. It is a large, ethnically diverse urban school system, which make the results relevant to other urban settings that face the challenge of educating at-risk students. NYC also provides a large sample of small high schools, which enables us to examine the relationship between size and performance *within* the sample of small schools, in addition to comparing small and large high schools.

In 2002, the New York State legislature granted Michael Bloomberg, the newly elected mayor of NYC, control of the NYC Public Schools. Mayor Bloomberg hired Joel Klein to be Chancellor, charging him with improving significantly the performance of NYC's one million plus public school students. One of Chancellor Klein's major initiatives was to establish new small high schools, replacing large dysfunctional ones. Over time, this strategy also involved providing high school students with a portfolio of schools from which they could choose to attend via an elaborate selection process modeled on the physician residency placement model.[4]

Chancellor Klein and his team succeeded in opening a very large number of new high schools.[5] Table 1 shows the numbers and characteristics of NYC high schools in 2007 an 2008 based on our sample of students and schools, which is described more fully below and includes the majority of non-special education high schools in these years. This sample is made up of two cohorts of NYC high school students who were scheduled for on-time (four-year) graduation in 2007 or 2008, were in the NYC public schools in 8th grade, attended non-special education schools, and attended schools with four grades by their graduation year. Note the large number of small schools (defined here as having 550 students) and student enrollment in small schools of just over 20% of our sample. In terms of the creation of *new* small schools, by 2008 there were 121 of them.[6] Fig. 1 shows the distribution of schools in our sample by size for the 2007 and 2008 cohorts compared to similarly constructed cohorts for 2001 and 2002 and clearly illustrates how much the size distribution was changed to favor small schools over these Klein-Bloomberg years.

Note that the NYCDOE set up a different creation process and regulatory environment for the new small schools established after 2002 compared to the old small schools that had been in existence before that time (Bloom et al., 2010). First, the new small schools came into being through a competitive application process in which school organizers proposed how they would institute academically rigorous curricula and partner with community organizations and not all applications were successful. Second, they were almost all supported by non-profit organizations, often New Visions for Public Schools (http://www.newvisions.org/), which

**Table 1**
Descriptive statistics of NYC HS students by school size category, 2008 and 2009 cohorts.

| | All | Small | Large |
|---|---|---|---|
| % Enrolled small schools | 20.26 | 100 | – |
| # Of schools | 291 | 169 | 122 |
| Distance to nearest small HS | 1.29 | 0.73 | 1.43 |
| Distance to nearest large HS | 0.64 | 0.59 | 0.66 |
| *Demographic characteristics* | | | |
| % Female | 53.36 | 56.94 | 52.45 |
| % Black | 37.25 | 42.90 | 35.82 |
| % Hispanic | 36.46 | 42.28 | 34.98 |
| % Asian | 14.46 | 7.63 | 16.20 |
| % White | 11.74 | 7.04 | 12.94 |
| % English is home language | 52.97 | 59.22 | 51.38 |
| % Overage | 16.61 | 17.83 | 16.31 |
| % Poor | 76.15 | 77.85 | 75.71 |
| % LEP | 3.25 | 3.02 | 3.31 |
| 8th grade Math *z*-score | 0.000 | −0.129 | 0.033 |
| 8th grade ELA *z*-score | 0.000 | −0.101 | 0.026 |
| *Outcomes* | | | |
| % Graduated | 65.76 | 72.04 | 64.16 |
| % Continued Enrollment | 21.74 | 18.62 | 22.53 |
| % Dropout | 9.84 | 7.56 | 10.43 |
| % GED | 2.66 | 1.78 | 2.88 |
| % Took Math Regents | 83.38 | 87.88 | 82.24 |
| % Score >55 | 96.97 | 96.51 | 97.10 |
| % Score >65 | 83.80 | 80.14 | 84.79 |
| % Took English Regents | 83.22 | 86.99 | 82.27 |
| % Score >55 | 96.01 | 95.49 | 96.16 |
| % Score >65 | 86.49 | 83.39 | 87.32 |
| Observations | 76,213 | 15,444 | 60,769 |

*Notes*: Small schools are those with enrollments of 550 or fewer students. Distance is calculated using Euclidian distances. Poverty is defined by eligibility for free lunch. ELA and math *z*-scores are taken from the statewide ELA and math exams given in 8th grade and are standardized to have a mean of zero and a standard deviation of one. Regents exams are New York State exams given to high school students. Schools that are considered small in one year, but large in the other are reported twice in the "number of schools" row.

[3] Put differently, the lottery based design of the Bloom study creates high internal validity, with potentially significant (although unknown) consequences for external validity and generalizability.

[4] The High School Application Processing System (HSAPS), was introduced in NYC in the 2003–2004 academic year. HSAPS requires all NYC public school 8th grade students to submit to the NYC Department of Education (NYCDOE) a list of up to 12 schools ranked in order of preference. The NYCDOE then uses a computerized process, modeled on the physician residency placement model, to assign students to their highest-ranked school whose admissions criteria they meet and where spaces are available. The matching takes into account students' preferences and schools' selection criteria and is supposed to be immune to gaming, thus encouraging students to list schools in order of their true choices (Abdulkadiroglu et al., 2009). Several high schools require entrance examinations, which students often take in the fall of their 8th grade year; students who qualify based on their test scores are offered places before the due date for rankings of other schools. In our empirical work, we perform robustness tests that exclude such schools to make sure that they are not driving results.

[5] According to Bloom et al. (2010), between 2002 and the time of their report (2010), NYC "... closed more than 20 underperforming public high schools, opened more than 200 new secondary schools... (p.12)" These numbers differ from ours in part because they run through 2010 while ours go through 2008 only.

[6] In this paper, a new school is one that had no graduates in the 2001 or 2002 cohorts but did have graduates in the 2007 or 2008 cohorts. A new school could have begun as early as 1999–2000, but generally new schools began with only a 9th grade class, adding one grade per year, thus not becoming full high schools until 2002–2003. The majority of schools labeled new began in 2002 or later (over 60% overall and over 62% for new small schools).

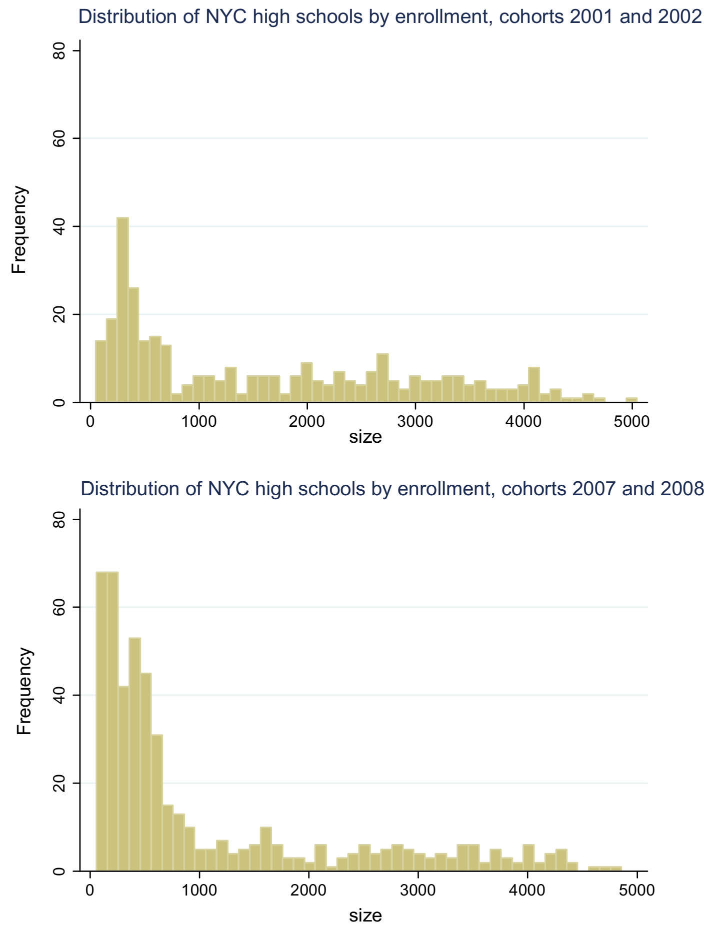


**Fig. 1.** Distribution of NYC high schools by enrollment.

were funded by the Gates and some other Foundations to monitor, aid and network them with other new small schools as they became established. Third, they were given some exemptions in their first years from serving some groups of special needs students and following all union rules on hiring teachers, although some of these exemptions were to disappear once they were fully functional. Last, their principals were often trained by an organization that was born through a NYCDOE effort to train new leaders who embraced accountability and empowerment by schools.[7] Later in this paper, we explore whether these differences were reflected in differences in quantitative characteristics of new small schools compared to existing small schools, such as their student composition, their teacher/student ratios, their per pupil expenditures, or their teacher characteristics.

A key concern of policy makers is the four year graduation rate. As Fig. 2 demonstrates there is a tremendous variation in the school-level graduation rate across the high schools in NYC. While some high schools achieve above 90% graduation rates, other schools see less than 40% of their students graduate. Thus, there is substantial heterogeneity in graduation rates across schools.

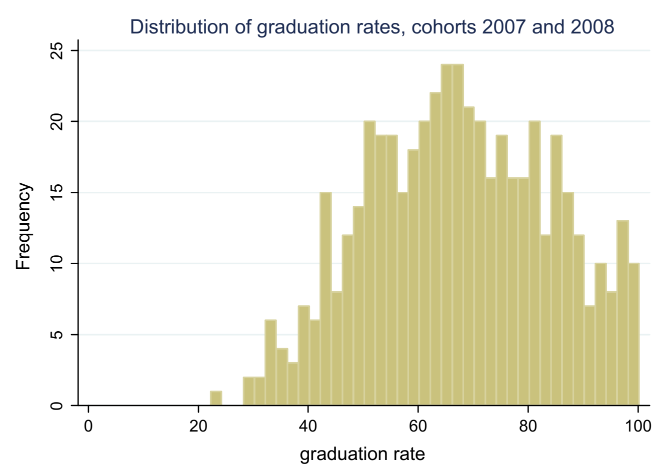


**Fig. 2.** Distribution of NYC high schools by graduation rates.

## 4. Data and measures

We use student- and school-level data from the NYC Department of Education (NYCDOE) administrative datasets. The student-level data are drawn from a census of NYC public high school students expected to graduate in either 2007 or 2008. Our sample of high schools includes those attended by the cohorts of students, with the exception of specialized program schools (such as last chance high schools or schools for pregnant mothers) or schools with predominantly full-time special education students. We exclude schools and students located on Staten Island because no small high schools exist there and students generally did not travel outside Staten Island to attend a small high school.

[7] See Corcoran et al. (2012) for more on the leadership academy.

The student data include student characteristics, such as socio-economic status, demographics and educational program participation, as well as a number of outcomes. Whether a student participated in the free lunch program in 8th grade serves as an indicator of poverty status,[8] and other demographic and education program variables include race/ethnicity, gender, whether or not a student was an English language learner, whether a student is overage for grade, and prior test scores in 8th grade reading and mathematics, which we convert to $z$ scores with mean zero and standard deviation one for each cohort in our sample. Also part of the student and school data are residence zip codes and school addresses, which enable us to calculate Euclidean distance, in miles, between home and schools.

Student high school outcomes include both graduation outcomes and high school test scores. Graduation is defined as earning a local, Regents, Honors or Advanced Regents diploma in four years. We exclude GEDs from the definition of graduation, but conduct robustness tests classifying obtaining a GED as graduation. The school we assign to students is the school in which they are enrolled in 9th grade. Although students are able to transfer among high schools, we assign them to their 9th grade school in order to obtain an "intent to treat" estimate.[9]

Students in NYC's public high schools must take statewide Regents examinations in a number of subject areas in order to receive a diploma. We focus on the Regents English and mathematics exams, as these were the first required of all students as part of the state's new graduation requirements that began in 1999–2000. We measure whether students take these exams, and conditional on taking them, whether their scores meet the cutoffs for various kinds of diplomas, cutoffs that changed over the time of our study. For the 2007 and 2008 cohorts, a minimal "local" diploma required a minimum score of 55 in any one of five core Regents areas,[10] and the higher level Regents diploma required a minimum score of 65 on all five of these exams. In addition, an Advanced Regents diploma was available to these later cohorts with scores of at least 65 on eight Regents exams. (See Appendix A for a more complete explanation of the Regents graduation requirements.)

School size is defined by the total number of students enrolled at each high school in a student's 9th grade year based on data from NYCDOE *School Based Expenditure Reports* and New York State *Annual School Reports*. Our primary definition of a small school is one with 550 or fewer students enrolled, although we analyze effects for larger sizes as well. Our reason for choosing this size is to be in the range of other studies and, in particular, to be consistent with a recent study in NYC, reviewed earlier (Bloom et al., 2010). A new school is one that had no graduates in the 2001 or 2002 cohorts but did have graduates in the 2007 or 2008 cohorts.

Table 1 displays the descriptive statistics for the combined 76,213 students in our study.[11] These students all began high school in 9th grade and are included in the graduation (or non-graduation) statistics. As discussed above, the number of small schools and enrollment in them increased substantially since 2001 and by the time of our study there were 169 of them. Comparing the descriptive statistics for small versus large high schools, we see that small schools overall have a more disadvantaged student population on a variety of measures. Small high schools have a higher proportion of their students eligible for free or reduced price lunches than do large high schools and more overage for grade. Small high schools students also score lower on 8th grade tests than students in large high schools.[12] The one exception to this pattern is that small school students have a higher proportion with English as a home language. Previewing our regression estimates described below, small schools have a higher overall graduation rate and a higher rate of students taking core mathematics and English Regents exams, but a lower or similar rate of achieving high scores (>=55 or >=65).

In summary, in both of the cohorts the composition of students in small and large high schools differs, with small high school students being generally less advantaged, making it essential to control for these characteristics in models of the effects of small school attendance on high school outcomes. For example, low-performing students may be more likely than other students to attend small schools, perhaps with the hope that less mainstream schools will turn their performance around or for some other unobserved reason related to performance. These students may experience gains in the small schools, yet continue to perform at lower levels than their large school counterparts. If that is the case, the average performance of students in these small schools compared to the rest of the schools will be lower, although the causal effect of small schools could still be positive.

## 5. Models and methods

### 5.1. Student performance model: common small school effect

Following the previous literature, we specify a stylized educational production function in which we model student outcomes as a function of observable variables capturing student sociodemographic and educational characteristics, performance on eighth grade English and math tests, and borough of residence, since in NYC these boroughs differ in their population income, education, and other demographics and could influence small school attendance and performance. Our basic model expresses student performance as follows:

$$Perf_{ijt} = \alpha_0 + \alpha_1 Small_{ijt} + ST'_{it}\alpha_2 + Test8'_{it}\alpha_3 + Borough'_{it}\alpha_4 + \varepsilon_{ijt} \tag{1}$$

where $Perf_{ijt}$ is a student outcome (such as earning a diploma within four years or taking a Regents examination) for student $i$ in school $j$ in year $t$. Since we use cohort data, as opposed to panel data, there is only one observation per student in each cohort dataset. $Small_{ijt}$ is an indicator that takes a value of 1 if, in year $t$, student $i$ attended a school $j$ that is small, as measured in this basic model, by a school that enrolls 550 or fewer students. $ST_{it}$ is a vector of student characteristics, including gender, race/ethnicity, free lunch status, English language proficiency, and overage for grade. $\mathrm{Test8}_{\mathrm{it}}$ is a vector of eighth grade reading and mathematics exam scores (each score, each score squared, and interacted), and $Borough_{it}$ is a set of indicators for the student's borough of residence (borough fixed effects). The $\alpha$'s are a set of intercepts and slopes that capture the impact of the corresponding variables on student performance, with $\alpha_1$, in particular, capturing the average difference in performance between students who attend small schools and students who do not, controlling for student characteristics and past performance. $\varepsilon_{ijt}$ represents the remaining variation. All standard errors are appropriately modified to reflect possible heteroskedasticity and clustering of students at the school level.

[8] We use an indicator of whether a student was eligible for free lunch in 8th grade rather than if she/he is eligible in high school, because the latter, unlike the former, is a notoriously poorly-reported variable.

[9] About 15–20% of the 2007–2008 cohort of students changed high schools.

[10] The five core Regents' exam areas are English, science, math, US history and government, and global studies.

[11] A table, available from authors and in the on-line appendix, contains definitions of each variable.

[12] By construction, the overall mean test score in our sample is zero (and the standard deviation is one).

### 5.2. Heterogeneity in treatment: new versus old small schools

Model (1) imposes two forms of homogeneity assumptions: (i) it assumes that the effect of small schools is the same for all students, and (ii) it assumes that the effect of small schools is the same for all types of small schools. Like much of the existing literature, we explore the student level heterogeneity (i) by estimating the effect of small schools for sub-groups defined by observable student characteristics. We focus here, however, on the school level heterogeneity by expanding (1) to allow for different effects of small school by the "vintage" of the school:

$$Perf_{ijt} = \alpha_0 + \beta_1 OldSmall_{ijt} + \beta_2 NewSmall_{ijt} + ST'_{it}\alpha_2 + Test8'_{it}\alpha_3 + Borough'_{it}\alpha_4 + \varepsilon_{ijt} \quad (2)$$

where $\beta_1$ is the effect of attending an old small school (defined as schools in operation prior to 2001–2002) and $\beta_2$ is the effect of attending a new small school (defined as schools that began operations in 2001–2002 or after). Since the new small schools differed in substantial ways from the old small schools (as discussed above), distinguishing between the effects of the different types of schools allows us to isolate whether school size is the key school characteristic for student performance or whether other features of the school are important.

### 5.3. Instrumental variable strategy

In order to overcome the possible selection of students into small schools on the basis of unobservable characteristics, we instrument for small school attendance with variables that plausibly affect school attendance but do not directly affect student outcomes except through small school attendance. As described earlier, these instruments are based on the minimum Euclidian distance between the nearest small school (either new or old small schools) or large school and the student's 8th grade residence zip code.[13] Note that we include borough fixed effects in the main specification so that we control for unobserved factors correlated with distance at the borough level.

#### 5.3.1. First stage of IV estimations

Table 2 reports linear probability models providing the relationship between small school attendance and various measures of distance. The full set of first stage estimates for our main results are available upon request. We also discuss below a series of robustness exercises to test whether our main results are sensitive to various configurations of the distance instruments.

In Table 2, the coefficients on the variable measuring distance from a small school and its square are statistically significant and plausible. Comparing Columns (1) and (2) we see that there is a concave relationship between distance and attending a small school with a negative linear term and positive quadratic term. For example, the probability on attending a small school as distance increases by one mile (from 0 miles away) decreases by 10%,[14] conditioning on student covariates and dummy variables for NYC boroughs. In contrast, distance to large school does not have a statistically significant effect of small school attendance in Column 2. As discussed below, in a series of robustness exercises, we also use distance to large schools as a control variable in the main specification, thus giving the distance to small school variables a relative distance interpretation. Our main results are robust to this instrument and control variable configuration. In results available on request, we show that distance to various vintages of small schools, new or old small schools, also has a statistically significant relationship with attendance at these particular school types, and that distance from large schools has either a statistically insignificant or quite small effect. Across our various IV specifications, the F-statistic for the total regression or for the excluded instruments is large, indicating that our distance instruments provide strong instruments for small school attendance (Staiger and Stock, 1997).[15]

**Table 2**
Probability of attending a small school and distance from residence, OLS regression results.

| Dependent variable: small school attendance | (1) | (2) |
|---|---|---|
| Distance to nearest small school | $-0.044^{***}$ | $-0.115^{***}$ |
| | (0.001) | (0.004) |
| Distance to nearest small school squared | | $0.015^{***}$ |
| | | (0.001) |
| Distance to nearest large school | $-0.009^{**}$ | $-0.012$ |
| | (0.004) | (0.010) |
| Distance to nearest large school squared | | 0.008 |
| | | (0.006) |
| Year 2008 | $0.028^{***}$ | $0.028^{***}$ |
| | (0.003) | (0.003) |
| Constant | $0.375^{***}$ | $0.411^{***}$ |
| | (0.006) | (0.007) |
| Observations | 76,213 | 76,213 |
| F – First stage excluded (2,76193) | $817.47^{***}$ | – |
| F – First stage excluded (4,76193) | – | $429.98^{***}$ |
| F – Total regression (19,76193) | $527.13^{***}$ | – |
| F – Total regression (21,76193) | – | $485.53^{***}$ |
| $R$-squared | 0.094 | 0.098 |

Robust standard errors in parentheses (*** $p<0.01$, ** $p<.05$, * $p<0.1$). All models control for gender, race/ethnicity, English proficiency, home language, overage for grade, poverty, performance on 8th grade standardized ELA and math exams, and residence borough. Poverty is measured by eligibility for free lunch. Test scores are measured as $z$-scores with a mean of zero and a standard deviation of one.

#### 5.3.2. Threats to validity

There are potentially a number of different threats to the validity of the distance based IV strategy. We discuss each of these in turn and when possible provide some evidence on their importance.

*5.3.2.1. Location of schools.* The first issue is that the location of the small schools could be correlated with unobservable characteristics, but there are institutional reasons to think that this might not be the case. Many small schools were co-located with other small schools in buildings vacated by very large schools. For example, while a few large schools are co-located, small schools are significantly more likely to be sharing space, with almost sixty percent of small schools sharing space for the 2008 cohort. Co-location is particularly prevalent among the new small schools: over two-thirds of new small schools share space with another school. These large, vacated schools had been in existence for a long period and were likely largely exogenously set with respect to the current distribution of unobservable student characteristics. Hence, we would expect that the location of most small schools is therefore related to the location of large schools and the long-run

[13] We use 8th grade home zip code for 2007 cohort. For 2008, only 7th grade home zip codes were available; home addresses were not available for any cohort. We calculate the distance between the centroid of each zip code and the school address. Distances are calculated using Stata 11 vincenty code. There are 170 student residence zip codes in our sample.

[14] Predicted change in probability of small school attendance from 0 to 1 mile away from school is $-0.115 * 1 + 0.015 * 1^2 = -0.1$.

[15] We use the quadratic in distance specification, as in column 2, Table 2, for the IV estimation.

historical factors which influenced the location of the large schools.[16]

While we condition on many observable student characteristics, whether school location is wholly unrelated to unobservable student characteristics is, to some extent, still an open question, one which we cannot resolve.[17] Supporting the validity of distance type instruments, however, Cullen et al. (2005) found in their study of Chicago public schools that distance to school was uncorrelated with a rich set of additional variables they collected, variables that would be unobservable in our administrative data.

Additionally, we performed a falsification test to further examine possible endogenous location of new small schools. We replicated our analyses for two earlier cohorts (2001 and 2002) that attended school *before* the new smaller high schools were opened. The coefficient from a regression of performance on 2007–2008 distances to new small school resulted in a small and insignificant coefficient on the distance variable. (Result available from authors.) This test would imply that new small schools were not placed strategically with respect to outcomes of students in earlier cohorts.

*5.3.2.2. Student mobility.* A related second issue is that a student's family might move to be closer to a particular school type, thereby creating a correlation between distance to school and unobservable student characteristics related to their perceived gains from these schools.[18]

There is little in the high school application process or system of preferences that creates a strong incentive to do so and there is little empirical evidence of such mobility. An examination of the residential moves made by 8th grade students in 2007 reveals that about nine percent changed zip code between 7th and 8th grade and those include both moves closer to and farther from small schools. We tested the relationship between these moves and school distances by regressing an indicator for student mobility on our set of student control variables and after-move distance to old and new small schools. (Result available from authors.) Controlling for student characteristics, we find no statistically significant difference in the distance to old or new small schools for individuals who move zip codes. This analysis suggests that for those students who do move, we cannot reject the hypothesis at the 10% level that these moves were unrelated to the location of old and new small high schools. While we find that distance matters for school choice, given the costs associated with traveling each day to a distantly located schools, we believe that other factors, including family and cultural ties to particular areas, job opportunities, and other neighborhood amenities may be far important factors in place of residence.

*5.3.2.3. Student entry and exit.* Another potential issue is that distance to various schooling options might have an impact on the set of students choosing to attend public high school, either causing some students to exit the NYC public system if they reside too far away from a desired high school option, or enter the NYC public system if they reside close to a desired option. Our student sample consists of only high school students who were also enrolled in NYC public middle schools. This sample restriction is necessary here in order for us to have available two of our key control variables, 8th grade mathematics and reading test scores. Thus our sample does not include students who entered or left the NYC public system between middle and high school.

It is important to emphasize that while distance to school types may be related to entry and exit decisions of students, this correlation does not affect the internal validity of the estimates we provide: *within* the sample of students who enroll in both NYC public middle and high school, distance to various school types is a plausible instrument. The issue of whether small school attendance would affect high school outcomes for the students who enter or exit the NYC system in the same way as we estimate for the population of students who remain in the system is fundamentally a question of external validity: Is the local average treatment effect (LATE) we identify representative for the whole population of high school students? The same extrapolation issue would of course apply to studies using lottery or discontinuities in admission criteria for identification as students may enter or leave the public system in response to the admission decisions (see Engberg et al. 2010 for a discussion of these issues).

*5.3.2.4. Spillovers.* Another possible threat to the validity of our IV strategy is that distance to a particular school may influence the outcomes of students not attending that school through various types of spillovers. There are at least several types of spillovers: (i) changes in student composition at other schools as students choose to attend various vintages of small schools, (ii) competitive pressures of small schools on other schools may induce principals in other schools to change their behavior in some way to make their school more attractive to students, and (iii) changes in school personnel as small schools attract teachers and principals.

While we cannot rule out these types of spillovers completely, and our estimates may reflect a combination of direct effects of small school attendance and spillover effects, we argue that the spillover effects are relatively small in magnitude. The primary reason is that the schools are small, hence the size of spillover effects, whether in changing peer groups at other schools, competitive pressure on other schools, or attracting school personnel would be limited. In addition, it is not clear in what direction these types of spillovers would bias the estimates. Spillover types (i) and (iii) would tend to make surrounding non-small schools worse, as the small schools attracted better students and better personnel, but the competitive pressure (ii) would tend to make other non-small schools better.

In regard to (ii) in particular, there are institutional reasons to believe this effect in particular might be limited. In NYC, this type of local, intra-district, competition is unlikely because the comparison groups by which school performance is judged are explicitly district-wide. NYC's accountability system compares schools that are similar in student body composition but very often geographically distant. Thus, the extent of competitive pressures from geographic distance may be considerably less relevant than across district competition.

We also conducted several empirical analyses to search for evidence of spillovers. To do this, we created a small school penetration variable by calculating the percentage of students in a zip code that attended a small high school. We then regressed features of large schools in each zip code against this penetration variable, conditional on school characteristics. We found small, insignificant coefficients on large school class size, teacher characteristics, and residuals of school graduation rates. (Result available from authors.) Thus small school penetration at the zip code level does

[16] Another concern may be that the large schools that closed during this period were purposefully chosen because they were among the worst performing schools in the district. This selection of large school closures would increase the average performance of the remaining large schools and therefore bias the effect of small new schools on performance downward and thus not change our overall qualitative conclusions.

[17] Although correlations between observable student characteristics and school location do not pose a threat to validity, we did estimate models examining the relationship between distance and average student level characteristics within a zip code. Results, available from authors, find distance is unrelated to the average observable characteristics of the zip code's student residents for all small schools and for new small schools, although there are some patterns for old small schools.

[18] In an analysis available from the authors, we examined the relationship between average observable student characteristics across zip codes and distance to schools by zip code and found few significant coefficients.

not seem to be related to class size, staffing or effectiveness of large schools.

Finally, even if the spillover effects are non-negligible, however, our results are still interpretable as the systemic effect of the availability of small schools. Of particular relevance to policy makers is how the real world impact of the introduction of alternative school types into an on-going school system. In other work, we find some support for the idea that small school reform works as systemic reform, at least in NYC, where our results suggest that graduation, Regents test taking, and Regents passing rates all improved significantly since 2001 in the large schools, in the small schools, in the continuously operating schools, and in the new schools overall in the city.

## 6. Results

### 6.1. Common small school model

In our most basic model, we estimate the effect of attending a small high school on earning a diploma in four years, using a linear probability model. Estimation using a probit model yields similar results (available from authors), but the linear probability specification is reported for ease of interpretation. In Table 3, Column (1) we estimate our basic model using OLS, including control variables for a number of student demographic characteristics (gender and race indicators), English proficiency and free lunch poverty status, and a quadratic function of 8th grade mathematics and reading scores (each score alone, squared, and their interaction). In addition, we include borough fixed effects. Controlling for these variables, the OLS estimated effect is positive and statistically significant from zero ($p < .01$). The coefficient estimate of 0.108 indicates that students are nearly 11% more likely to graduate if they attend a small rather than a large high school.

These results, however, are potentially suspect because, as discussed earlier, students may decide to attend small schools based on unobserved characteristics that could also predispose them to be more or less likely to earn a diploma, and such selection bias could affect the size and the sign of the coefficient on the small school indicator. To address this source of selection bias, we employ an instrumental variable approach, using two stage least squares (2SLS), with instruments for small school attendance based on distance from the students' homes to the nearest small and large high schools. The first stage estimates are available from the authors (see the discussion of instrument strength and validity above).

Table 3 Column (2) displays the second stage IV estimates for the effect of small school attendance on earning a diploma in 4 years. As with the OLS results, all of these models include the full set of student covariates and borough fixed effects. The IV results differ dramatically from the OLS estimates as we now obtain a small and not statistically significant (at $p < .10$) effect of small school attendance.

If we were to end the analysis here, we would draw the following conclusions. Although on observable characteristics more disadvantaged students attend small schools in NYC, there seems to be positive selection on unobservable variables into the small schools given the positive OLS estimates of small school attendance and the negative or insignificant from zero IV estimates. After correcting estimates for selection on unobservables, using strong instruments related to distance from residence to school, the coefficient on the small school variable changes signs and significance, from positive in the OLS results to insignificant in the IV results. The unobservables that are affecting selection into small high schools are not knowable using our data, but they may plausibly be related to motivation and/or parental involvement. This result implies that it is not school size but rather selected students that make the OLS estimates of small school effects positive. Stopping here, however, might not present the entire story since small schools are not all created equal, outcomes other than earning a diploma can be important, and we have only explored one definition of size. We turn to these issues next.

**Table 3**
Probability of graduating in four years, OLS and IV regression results.

| Dependent variable: graduated | (1) OLS | (2) IV | (3) OLS | (4) IV |
|---|---|---|---|---|
| Small | 0.108*** | −0.011 | | |
| | (0.011) | (0.058) | | |
| Old small | | | 0.072*** | −0.556*** |
| | | | (0.016) | (0.167) |
| New small | | | 0.125*** | 0.175*** |
| | | | (0.012) | (0.084) |
| Year 2008 | 0.039*** | 0.043*** | 0.037*** | 0.026*** |
| | (0.005) | (0.006) | (0.005) | (0.008) |
| Constant | 0.664*** | 0.704*** | 0.663*** | 0.710*** |
| | (0.012) | (0.023) | (0.012) | (0.029) |
| Observations | 76,213 | 76,213 | 76,213 | 76,213 |
| $R$-squared | 0.236 | 0.227 | 0.237 | 0.131 |
| F (old sm = new sm) | | | 10.52*** | 13.84*** |
| F (old sm = new sm = 0) | | | 58.42*** | 6.92*** |

Robust standard errors, adjusted for within-school clusters, in parentheses (*** p<0.01, ** p<0.05, * $p < 0.1$). All models control for gender, race/ethnicity, English proficiency, home language, overage for grade, poverty, performance on 8th grade standardized ELA and math exams, and residence borough. Poverty is measured by eligibility for free lunch. Test scores are measured as $z$-scores with a mean of zero and a standard deviation of one. Graduated is defined as earning a local, Regents, or Regents Honors diploma, as defined by the New York State Department of Education.

### 6.2. New versus old small schools

We next explore whether some small schools are different than others, and specifically whether small schools that were newly created, with the extra supports and rigorous application processes (and perhaps other differences), are more effective than the earlier generation of small schools. We examine the descriptive statistics on differences between the old small and new small schools in Table 4. The new small schools have smaller enrollment and more advantaged students than the old small schools. The students at the new small schools have similar levels of poverty (measured by free lunch eligibility) as the students at the old small schools, but the new small school students have a higher fraction of Asian and higher 8th grade English and mathematics scores.

To examine the potentially differential effectiveness of these different types of schools, we re-estimate the OLS and IV models separately for new and old small high schools. In Table 3, the OLS estimates in Column (3) indicate that both new and old small schools have statistically significant positive effects on graduation, relative to large schools, but that the new small schools have a higher positive effect (0.125 versus 0.072). The $F$-test statistics indicate that we reject ($p < .01$) the joint null hypotheses that the effect of the small schools is the same and that both have zero effect.

Because the same endogenous school choice issue could exist for our two different school types as with the single small school type, we next turn to IV estimation. In column (4) of Table 3, we compute 2SLS estimates and instrument for both endogenous old small and new small high school attendance.[19] Since this new specification includes multiple endogenous school choice variables,

[19] The two first stage results are available from the authors.

**Table 4**
Descriptive statistics of New York City high school students by new small and old small high school.

| | New small | Old small |
|---|---|---|
| % Enrolled | 13.83 | 6.43 |
| # Of Schools | 121 | 48 |
| Average size | 245 | 420 |
| *Demographic characteristics* | | |
| % Female | 56.32 | 58.27 |
| % Black | 43.20 | 42.24 |
| % Hispanic | 40.62 | 45.83 |
| % Asian | 8.67 | 5.38 |
| % White | 7.33 | 6.42 |
| % Overage | 17.25 | 19.07 |
| % LEP | 2.98 | 3.10 |
| % English is home language | 59.70 | 58.19 |
| % Poor | 77.48 | 78.65 |
| 8th grade Math *z*-score | −0.084 | −0.224 |
| 8th grade ELA *z*-score | −0.061 | −0.187 |
| *Outcomes* | | |
| % Graduated | 74.83 | 66.04 |
| % Continued Enrollment | 16.65 | 22.86 |
| % Dropout | 6.96 | 8.83 |
| % GED | 1.56 | 2.26 |

*Notes*: Small schools are those with enrollments of 550 or fewer students. Poverty is defined by eligibility for free lunch. ELA and math *z*-scores are taken from the statewide ELA and math exams given in 8th grade and are standardized to have a mean of zero and a standard deviation of one. New schools are schools with a graduating class in 2003; old schools are schools with graduating classes prior to 2003.

identification of the IV estimator requires multiple instruments. Our instruments are based on a set of distance based instrumental variables: student residence distance to old small schools, distance to new small schools, distance to old large schools, and distance new large schools. The first stage estimates indicate a negative and statistically significant relationship between distance from student residence to old small schools and old small school attendance and between distance to new small schools and new small school attendance. We explore below various other instrumental variable configurations and find our main estimate results are robust to alternative IV strategies.

The results in Table 3 suggest that the OLS result is driven by selection on unobservable student characteristics as the IV estimates are considerably different from the corresponding OLS estimates. The IV results show that attending an old small high school had a negative effect on performance ($p < .01$), while attending a new small school had a positive effect ($p < .01$). The negative estimate for old small schools indicates that older generation of small schools were substantially worse than the large high schools (the omitted category of schools). Note that while this negative effect is quite sizable (−0.556 relative to large schools), the robustness analysis below suggests that the negative effect may be driven by a few outlier schools since we estimate a smaller negative effect with different cutoff values defining "small" schools (and thus large schools). The positive estimate for new small schools indicates a 17.5% gain in the probability of graduating from high school due to attending a new small school. The estimates thus reveal a ranking in the causal effectiveness of schools in producing rates of earned diplomas in four years: old small schools are the worst, large high schools are better than old small schools, and the new small schools are better than large high schools.[20]

[20] In previous analyses, we have found a "newness" effect for large schools in OLS but not IV estimations, but the sample of new large schools (15) and students (2000) is too small to be relied upon. Perhaps more importantly for policy purposes, we focus on new small schools because the current reform initiatives have created many more new small schools than new large schools.

### 6.3. Robustness

In results available on request, we have examined the robustness of our results to several changes in model specification, including dropping selective high schools, changing the definition of small high schools, and using various alternative instruments constructed from functions of the original instruments. Overall our main estimates are robust to these alternative IV models. Across all of these alternative IV models, the causal effect of new small attendance ranges from 0.155 to 0.199, with all of the estimates precisely estimated at least at the 10% level. The negative effect of old small school attendance shows more sensitivity to the IV estimator choice, with the negative effect relative to large schools ranging from −0.252 to −0.517, with the −0.252 effect estimated imprecisely. Across all models, however, the *F*-test statistic on the joint hypotheses that small and old schools have the same effect on graduation and the *F*-test statistic that old and new small schools jointly have a zero effect are generally large and we can reject these hypotheses at the 1% level.

### 6.4. Other outcomes

In Table 5 we explore a number of different high school outcomes using our main IV specification, instrumenting for endogenous old small and new small school attendance. The omitted category is large high schools, so all effect estimates are gains or losses relative to attending a large high school. In Column (1) we examine how attendance at old and new small schools affects the probability of earning a GED degree instead of a regular or Regents diploma (our "graduate" outcome) or not earning any degree (high school drop-out). The estimates reveal that attending a new small school lowers the probability of earning a GED by nearly 6%. This can be considered a desirable outcome combined with the positive 17% effect of new small school attendance on the probability of graduating. The estimates suggest that new small schools cause some students who would have otherwise earned a GED to instead graduate from high school. In contrast, the effect of old small school attendance has a small and imprecise effect on earning a GED.

Earning a diploma in four years is a critical outcome, but the quality of the high school degree is important as well. Moreover, most states, New York included, are making passage of content examinations a requirement for graduation. To ascertain whether attending a small school has an effect on either the number of examinations taken or the scores achieved, we estimate IV results for two critical New York State Regents examinations – English and first level mathematics.

Table 5 provides the IV estimates for a range of these examinations. New small high schools appear to be considerably more effective than the old small schools on all of these measures. Mirroring the effects on graduation, new small school attendance also has a statistically significant effect on attempting Regents mathematics and English examinations. Our IV estimates indicate that attending a new small school relative to a large school increases the probability of taking the Regents English by 14.4% and the Regents mathematics by 16.4%.

Perhaps because of the large effect on test taking, attendance at a new small schools is estimated to cause a reduction in average English scores and the likelihood of obtaining passing scores (for those who take the exam). New small schools attendance is estimated to reduce the student's English score by 0.4 standard deviations and reduce the probability of obtaining a score of 55 or higher than 55 score by 4.5% and a score of 65 higher than 65 score by 19.8%. While these results are seemingly in contradiction with the positive graduation effects, they may reflect the much higher proportion of students taking the English exam and that the

**Table 5**
Other education outcomes, IV regression results.

| | (1) GED | (2) Took English | (3) English ⩾55 | (4) English ⩾65 | (5) English *z*-score | (6) Took Math | (7) Math ⩾55 | (8) Math ⩾65 | (9) Math *z*-score |
|---|---|---|---|---|---|---|---|---|---|
| Old small | −0.013 | −0.296*** | −0.152*** | −0.430*** | −0.807** | −0.251** | −0.086** | −0.431*** | −0.517*** |
| | (0.032) | (0.109) | (0.051) | (0.126) | (0.346) | (0.114) | (0.037) | (0.125) | (0.196) |
| New small | −0.058*** | 0.144** | −0.045** | −0.198*** | −0.402** | 0.164*** | −0.016 | −0.066 | −0.021 |
| | (0.021) | (0.056) | (0.020) | (0.059) | (0.193) | (0.061) | (0.015) | (0.054) | (0.100) |
| Year 2008 | 0.002 | 0.034*** | −0.003 | 0.022*** | −0.131*** | 0.040*** | −0.008*** | 0.009 | 0.062*** |
| | (0.002) | (0.006) | (0.002) | (0.007) | (0.041) | (0.007) | (0.002) | (0.006) | (0.013) |
| Constant | 0.048*** | 0.847*** | 0.835*** | 1.006*** | 0.981*** | 0.622*** | 1.005*** | 0.934*** | 0.228*** |
| | (0.007) | (0.021) | (0.022) | (0.009) | (0.026) | (0.100) | (0.006) | (0.023) | (0.038) |
| Observations | 76,213 | 76,213 | 76,213 | 63,427 | 63,427 | 63,427 | 63,547 | 63,547 | 63,547 |
| *R*-squared | 0.001 | 0.080 | 0.089 | 0.068 | 0.089 | 0.256 | 0.084 | 0.162 | 0.480 |
| F (old = new) | 1.32 | 12.45*** | 3.51* | 2.65 | 1.36 | 9.71*** | 2.68 | 7.50*** | 5.12** |
| F (old = new = 0) | 3.84** | 6.72*** | 7.98*** | 12.15*** | 3.84** | 5.65*** | 3.67** | 6.41*** | 3.48** |

Robust standard errors, adjusted for within-school clusters, in parentheses. All models control for gender, race/ethnicity, English proficiency, home language, overage for grade, poverty, performance on 8th grade standardized ELA and math exams, and residence borough. Poverty is measured by eligibility for free lunch. Test scores are measured as *z*-scores with a mean of zero and a standard deviation of one. Graduated is defined as earning a local, Regents, or Regents Honors diploma, as defined by the New York State Department of Education. English and Math are two of the many NYS Regents examinations given three times annually. As students can re-take these exams, we use their most recent score. Columns (6) and (10) use student scores standardized to have a mean of zero and a standard deviation of one for all student taking the exam on that particular date.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

marginal test taker had lower ability.[21] Also recall that for our sample years, the minimum local diploma requires that only one Regents Examination be passed with a score of 55%.

In addition, Table 5 reveals that the new small schools still have much better performance than old small schools in terms of test taking and performance, although we cannot reject the hypothesis at the 10% level that the effect of new and old small on average English scores is the same.

In contrast to the English score results, we estimate that new small school attendance has a small negative and imprecise effect on math performance. Thus, while there is also a large and positive effect on test taking in mathematics as with English, there is no significant negative effect of new small school attendance on math performance. For old small schools, however, we estimate that attendance would not only reduce the probability of taking the mathematics exam but also substantially reduce the average score and probability of obtaining a passing score. Attendance at old small schools is estimated to reduce math scores by 0.517 standard deviations and the probability of earning a score of 65 or greater than a 55 score is reduced by 8.6% and greater than a 65 score by 43%.

### *6.5. Student sub-group analysis*

The literature on small high schools hypothesizes that particular benefits accrue to at-risk subgroups such as black, Hispanic, poor and/or male students. In Table 6, we present empirical evidence on graduation rates for subgroups attending NYC high schools. Based on IV regression results, the new small schools perform better than the old small schools for all subgroups and better than large high schools for both girls and boys, Asian, Hispanic, white, and non-poor students. For black and poor students, attending new small and large high schools is equivalent in terms of graduating in four years. Thus, students attending old small high schools clearly perform worse than those attending other size schools, while students attending new small high schools either perform better or equivalent to those attending large high schools.

## 7. Discussion: why does “newness” matter?

If only the *new* small high schools are more effective than large high schools, it is important to explore how the characteristics across types of schools might differ. We already noted the kinds of process differences that NYCDOE identified for new small schools, but are there also differences in the characteristics of the student bodies, in spending per pupil, or in teacher characteristics? Table 7 displays OLS results for a series of school-level variables and allows us to begin to answer this question.

In terms of student racial composition and in comparison to old small schools, both the new and old small schools enroll lower percentages of Asian students. They do not differ from each other or large schools in terms of the percentages of enrolled black or Hispanic students, however. New small schools enroll fewer special education students compared to both large schools and old small schools and fewer limited English proficient (LEP) students compared to large schools (but not compared to old small schools). The percentage of poor students at new small schools is higher than at old small and large schools.

Strikingly, resources for small high schools differ, with both new and old small schools receiving more funding for direct (school-level) resources and having lower class sizes (pupil-teacher ratios) compared to large schools. These differences are substantial (between $1700 and almost $2900 per pupil and around 2.8 fewer students per teacher). Small schools also have more inexperienced teachers and fewer with MA degrees (with new small significantly different from old small and large), which implies that the extra funding is financing lower class sizes and perhaps other services as well, but with less experienced and educated teachers. Finally, only old small schools differ from large schools in terms of principals trained at the newly created Leadership Academy set up to provide good principals to school.

Thus, the new small schools receive significantly more resources, work with less trained teachers, serve more poor but fewer special education and LEP students than large schools and sometimes than old small schools as well. These differences along

[21] Because test taking is voluntary and likely related to unobservable variables that also affect high school choice, considering the test score outcomes introduces a “double endogeneity” problem with both test taking and school choice being endogenous. While we have a plausible IV strategy for identifying school effects overall, we do not have a credible instrument for test taking. Understanding the important relationships between test taking and school attendance choices is an important area for future research.

**Table 6**
Subgroup analyses, IV regression results.

| Dependent variable: grad. | (1) Female | (2) Male | (3) Asian | (4) Black | (5) Hispanic | (6) White | (7) Poor | (8) Non-poor |
|---|---|---|---|---|---|---|---|---|
| Old small | −0.647*** | −0.435** | −0.281 | −0.563** | −0.337** | −0.092 | −0.462*** | −0.437** |
| | (0.182) | (0.192) | (0.284) | (0.243) | (0.170) | (0.251) | (0.156) | (0.218) |
| New small | 0.177* | 0.162* | 0.241*** | 0.031 | 0.316*** | 0.225* | 0.125 | 0.228** |
| | (0.091) | (0.098) | (0.088) | (0.112) | (0.122) | (0.125) | (0.102) | (0.093) |
| Year 2008 | 0.023** | 0.031*** | 0.000 | 0.037*** | 0.031** | 0.033*** | 0.030*** | 0.026*** |
| | (0.009) | (0.009) | (0.007) | (0.012) | (0.012) | (0.009) | (0.010) | (0.009) |
| Constant | 0.798*** | 0.697*** | 0.756*** | 0.874*** | 0.633*** | 0.725*** | 0.643*** | 0.697*** |
| | (0.031) | (0.033) | (0.034) | (0.040) | (0.037) | (0.037) | (0.027) | (0.037) |
| Observations | 40,669 | 35,544 | 11,024 | 28,390 | 27,849 | 8,950 | 58,034 | 18,179 |
| *R*-squared | 0.060 | 0.180 | 0.168 | 0.107 | 0.115 | 0.221 | 0.143 | 0.114 |
| F (old = new) | 15.02*** | 6.61** | 3.69* | 4.92** | 7.97*** | 1.05 | 8.22*** | 8.11*** |
| F (old = new=0) | 7.52*** | 3.32** | 5.53*** | 2.71* | 4.34** | 1.63 | 4.53** | 5.20*** |

Robust standard errors, adjusted for within-school clusters, in parentheses. All models control for gender, race/ethnicity, English proficiency, home language, overage for grade, poverty, performance on 8th grade standardized ELA and math exams, and residence borough. Poverty is measured by eligibility for free lunch. Test scores are measured as *z*-scores with a mean of zero and a standard deviation of one. Graduated is defined as earning a local, Regents, or Regents Honors diploma, as defined by the New York State Department of Education.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

**Table 7**
School characteristics, OLS regression results.

| | (1) % Asian | (2) % Black | (3) % Hispanic | (4) % Spec ed | (5) % LEP | (6) % Poor | (7) $-direct | (8) Pupil-teacher | (9) % tch < 2 yrs | (10) % MA+ | (11) tchr experience | (12) LA principal |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Old small | −7.777*** | 3.464 | 8.078** | 1.036 | −4.561*** | 6.446 | 1,719.78*** | −2.892*** | 7.611*** | −3.795** | −11.338*** | 0.141** |
| | (1.425) | (4.060) | (3.747) | (0.998) | (1.061) | (4.051) | (312.083) | (0.418) | (2.692) | (1.800) | (2.235) | (0.069) |
| New small | −5.096*** | 6.331* | 4.877* | −3.602*** | −2.551** | 10.613*** | 2,894.71*** | −2.815*** | 55.070*** | −19.768*** | −33.171*** | 0.040 |
| | (1.691) | (3.228) | (2.860) | (0.719) | (1.232) | (2.758) | (273.788) | (0.340) | (1.934) | (3.336) | (2.426) | (0.040) |
| Constant | 11.673*** | 39.092*** | 37.002*** | 10.842*** | 10.754*** | 57.377*** | 9,568.04*** | 17.544*** | 40.310*** | 76.422*** | 55.604*** | 0.086*** |
| | (1.228) | (2.485) | (2.166) | (0.490) | (0.825) | (2.141) | (169.018) | (0.227) | (1.322) | (0.662) | (1.024) | (0.026) |
| Observations | 278 | 278 | 278 | 278 | 278 | 278 | 276 | 276 | 255 | 255 | 255 | 279 |
| R-squared | 0.061 | 0.014 | 0.020 | 0.112 | 0.036 | 0.050 | 0.298 | 0.232 | 0.765 | 0.164 | 0.467 | 0.021 |
| F (old = new) | 3.84* | 0.56 | 0.80 | 20.81*** | 3.15* | 1.17 | 11.98*** | 0.03 | 300.69*** | 18.91*** | 54.27*** | 2.06 |
| F (old = new = 0) | 15.02*** | 1.93 | 2.68* | 16.69*** | 9.26*** | 7.40*** | 58.09*** | 43.05*** | 431.12*** | 18.97*** | 96.51*** | 2.23 |

Robust standard errors, adjusted for within-school clusters, in parentheses. One school is missing all information in 2004 and 2005. Two additional schools are missing spending and pupil-teacher information in 2004 and 2005. Twenty-four schools are missing information about teachers. Column (9) reports the results for the percent of teachers who have been teaching at that school for less than two years. The percent of teachers with a masters degree or higher is shown in column (10). Teacher experience, in column (11) is defined as the percent of teachers at the school who have 5+ years of teaching experience. Column (12) reports the results for whether the school was ever run by a Leadership Academy principal.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

with the processes the NYCDOE uses to choose among applicants proposing to form a new small school may set the stage for their success.[22]

## 8. Conclusions

District school reforms generally take several years to implement, and small high school reforms, which have been initiated in many major urban districts over the past two decades, are no different. In this paper, we use data on two cohorts of NYC high school students to estimate the effect of attending small high schools. A particularly important feature of this evaluation is that many small high schools existed from previous waves of creation in the 1990s, allowing us to test for the effect of smallness versus newness in school size. In order to ascertain if the size of the school is a *cause* of any changes in outcomes rather than just a product of the particular students who chose to attend these schools, we use instrumental variable estimators with plausibly exogenous distance from residence to school to correct for self-selection into schools.

We find first that selection is likely to be a complicating factor as we find very different OLS and IV effects of small school attendance on student performance. Decomposing small high schools into their earlier wave (fully created before 2002) and a later wave, we find that the later, new small schools are the ones that are effective in terms of graduation as well as taking math and English examinations. Thus the findings on size itself are ambiguous. Attending *new* small high schools can benefit students, but *old* small schools are not, on average, better. In fact, as a caution to associating school size automatically with improvements in stu-

[22] Some new large schools were formed after 2002, but these were few (15) and few students attended (only 2000 out of our sample of over 76,000), making estimation of their effects difficult. Nonetheless, in models where we include these schools with the old small and the new small, the OLS coefficient on new large is positive and significant but the coefficient from the IV estimation is insignificant. If the sample of new large schools were larger, this result would imply that neither newness or small size is the driving characteristic, but rather new and small together are necessary.

dent outcomes, we find that the older generation of small schools was actually associated with *worse* outcomes than the newer small schools or even some types of larger high schools.

While the new and old small schools differ in a variety of ways, including greater resources per pupil, the most important differences may lie in NYCDOE institutional policies that govern their creation and practices such as the application process they must go through, the monetary and networking support received from non-profit organizations, the loosening of union rules in hiring teachers, and the temporary suspension of requirements to serve all special needs students. These differences raise a host of questions for policymakers in NYC and for those hoping to replicate the success in other districts.

Perhaps most obvious and important are these: are these supports sustainable for the new small schools or can the schools be weaned? Can the successes of the new small schools be replicated in old small schools? Or is it the enthusiasm of teachers, leaders and staff energized by participating in building something "new" (or the ability of a new leader to choose staff to meet a vision) that creates the effect, meaning "growing old" is the problem? Does the success hinge on the increased funding? If size itself is not the defining characteristic, then just reducing school size will not produce the same results. While the notion that "small size is not enough" seems (with the benefit of hindsight) to have significant intuitive appeal, urban districts around the country have jumped on the small school reform bandwagon, replacing large, comprehensive schools with a myriad of small schools and schools-within-schools (with a wide range of resources, supports and characteristics) in the hopes of bolstering student performance. The evidence from New York City suggests that the success of these efforts will depend significantly on how those new schools are created and supported.

## Acknowledgments

We thank Elizabeth Debraggio and Jennifer Sallman for patient and excellent research assistance and the Institute for Education Sciences, US Department of Education, Grant Number R305A080522 for financial support. All analyses and conclusions are ours alone.

## Appendix A. Regents examinations

The regents examinations are a series of tests, aligned with New York State's Learning Standards, which New York students must pass in order to receive high school diplomas. They are designed and administered under the authority of the Board of Regents of the University of The State Of New York (the State governing body for K-16 education) and prepared by teacher examination committees and testing specialists. Examination scores range from 0% to 100%.

To earn a Regents high school diploma, New York students need to obtain appropriate credits in a number of specific subjects by passing year-long or half-year courses, after which they must pass a Regents examination in that subject area. This expectation is in addition to passing the courses themselves, the passing grade of which is based on an individual teacher's or school's own tests and class work. Starting with the cohort entering grade 9 in 2001, and thus including our own cohorts, to receive a Regents high school diploma students need to score a 65 or above in the following five content areas: Integrated Algebra (or Math A), Global History and Geography, U.S. History and Government, Comprehensive English, and any one science area.[23] To earn an Advanced Regents diploma, students take additional credits in a foreign language, pass an additional Regents exam in science (at least one course in life science and one in physical science), and pass a second Regents exam in math. Students in our cohorts also were allowed to graduate with local (not Regents) diplomas, which required passing any one of five Regents examinations with a score of at least 55%.[24]

The math exams offered for the cohorts in our study are Math A and Math B. Topics tested by the Math A Regents exam include equations and inequalities, probability and statistics and geometry. Math B, which is optional, is taken after the student has passed Math A. Topics that can be tested include concepts from trigonometry and advanced algebra, as well as some pre-calculus and calculus.[25]

New York State's science core curricula include Living Environment, Physical Setting/Earth Science, Physical Setting/Chemistry, and Physical Setting/Physics. All students entering grade 9 in 2001 must earn three units of credit in science although they must pass only one Regents examination in science to obtain a Regents diploma. The three science credits must be comprised of commencement-level science courses, including one course from the Physical Setting (physical science) and one course from the Living Environment (life science). The third may be from either life sciences or physical sciences. All commencement-level science courses, including specialized courses, must include laboratory activities.[26]

## Appendix B. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jue.2013.03.008.

## References

Abdulkadiroglu, A. et al., 2009. Strategy-proofness versus efficiency in matching with indifferences: Redesigning the NYC high school math. American Economic Review 99 (5), 1954–1978.

Armour, D.L., Peiser, B.M., 1998. Interdistrict choice in Massachusetts. In: Peterson, P.E., Hassel, B.C. (Eds.), Learning From School Choice. Brookings Institution Press, Washington, DC, pp. 157–186.

Barker, R.G., Gump, P.V., 1964. Big School. Stanford University Press, Small School, Stanford, CA.

Barrow, L., Claessens, A., Whitmore Schanzenbach, D., 2010. The impact of small schools in Chicago: Assessing the effectiveness of Chicago's small high school initiative, February. Working paper.

Bloom, H., et al., 2010. Transforming the High School Experience. Report June, MDRC, New York.

Booker, K. et al., 2011. The effects of charter high schools on educational attainment. Journal of Labor Economics 29 (2), 377–415.

Burgess, S., Briggs, A., 2010. School assignment, school choice and social mobility. Economics of Education Review 29 (4), 639–649.

Card, D., 1995. Using geographic variation in college proximity to estimate the return to schooling. In: Christofides, L.N., Grant, E.K., Swidinsky, R. (Eds.), Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp. University of Toronto Press, Toronto.

Conant, J.B., 1959. The American High School Today. McGraw-Hill, New York.

Corcoran, S.P., et al., 2012 forthcoming. Training your own: The impact of New York City's aspiring principals program on achievement. Educational Evaluation and Policy Analysis.

Cullen, J.B. et al., 2005. The impact of school choice on student outcomes: An analysis of Chicago public schools. Journal of Public Economics 89, 729–760.

Currie, J., Morretti, E., 2003. Mother's education and the intergenerational transmission of human capital: evidence from college openings. Quarterly Journal of Economics 118 (4), 2003.

[23] http://schools.nyc.gov/NR/rdonlyres/53FADF0D-D784-435E-8675-90E20624DAE1/0/2011GenEd.pdf.

[24] The local diploma option was gradually made more difficult (more examinations and higher grades required) and was finally removed as an option for students entering 9th grade in 2008 and later.

[25] The New York State Math A and Math B Regents Examinations are no longer a part of the High School Mathematics curriculum. The last administration of the Regents Examination in Mathematics A was January 2009 and the last administration of the Regents Examination in Mathematics B was June 2010. These exams were replaced by three exams: Integrated Algebra I, Geometry and Algebra II, and Trigonometry.

[26] For more information, see http://www.p12.nysed.gov/part100/pages/1005.html#regentsdiploma.

Engberg, J., et al., 2010. Evaluating education programs that have lotteried admission and selective attrition. Working paper.
Fowler, Jr., W.J., 1992. What do we know about school size? – What should we know? In: Paper presented at the annual meeting of the American Educational Research Association. San Francisco, CA.
Fowler Jr., W.J., Walberg, H.J., 1991. School size, characteristics, and outcomes. Educational Evaluation and Policy Analysis 13, 189–202.
Gootman, E., 2005. Gates donates $28.5 million for support of small schools. The New York Times (February 16).
Greene, J.P. et al., 1998. Lessons from the Cleveland scholarship program. In: Peterson, P.E., Hassel, B.C. (Eds.), Learning From School Choice. Brookings Institution Press, Washington DC.
Haller, E.J. et al., 1990. School size and program comprehensiveness: evidence from high school and beyond. Educational Evaluation and Policy Analysis 12, 109–120.
Hastings, J.S., et al., 2006. Parental Preferences and School competition: Evidence from a Public School Choice Program. NBER Working Paper, 11805.
Heckman, J., Navarro-Lozano, S., 2004. Using matching, instrumental variables and control functions to estimate economic choice models. Review of Economics and Statistics 86, 30–57.
Howley, C., et al., 2000. Research about School Size and School Performance in Impoverished Communities. ERIC Clearinghouse on Rural Education and Small Schools. ERIC Document Reproduction Service No. ED448968, Charleston, WV.
Hoyt, W.H., Lee, K., 1998. Educational vouchers, welfare effects and voting. Journal of Public Economics 69, 211–228.
Ladd, H., 2003. Comment on Caroline M. Hoxby: school choice and school competition: evidence from the United States. Swedish Economic Policy Review 10, 67–76.
Lee, V.E., Smith, J.B., 1995. Effects of high school restructuring and size on early gains in achievement and engagement. Sociology of Education 68, 241–270.
Lee, V.E., Smith, J.B., 1997. High school size: which works best and for whom? Educational Evaluation and Policy Analysis 19, 205–227.
Lindsay, P., 1982. The effect of high school size on student participation, satisfaction, and attendance. Educational Evaluation and Policy Analysis 4, 57–65.
McMillen, B.J., 2004. School size, achievement, and achievement gaps. Educational Evaluation and Policy Analysis 12 (58).
National Center for Education Statistics, 2010. The Nation's Report Card: Grade 12 Reading and Mathematics 2009 National and Pilot State Results. NCES 2011–455. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington DC.
Page, L. et al., 2002. National Evaluation of Smaller Learning Communities: Literature Review. Abt Associates, Inc., Cambridge, MA.
Saporito, S., Lareau, A., 1999. School selection as a process: the multiple dimensions of race in framing educational choice. Social Problems 46, 418–439.
Schneider, M., Buckley, J., 2002. What do parents want from schools? Evidence from the internet. Educational Evaluation and Policy Analysis 24 (2), 133–144.
Schneider, M. et al., 1998. School choice and culture wars in the classroom: what different parents seek from education. Social Science Quarterly 79, 489–501.
Schneider, B. et al., 2007. Is small really better? Testing some assumptions about high school size. In: Loveless, T., Hess, F.M. (Eds.), Brookings Papers on Education Policy 2006/2007. The Brookings Institution, Washington DC.
Stiefel, L. et al., 2000. High school size: effects on budgets and performance in New York City. Educational Evaluation and Policy Analysis 22, 27–39.
Stiefel, L. et al., 2009. Mission matters: the cost of small high schools revisited. Economics of Education Review 28 (5), 585–599.
Staiger, D., Stock, J., 1997. Instrumental variables regression with weak instruments. Econometrica 65 (3), 557–586.
U.S. Department of Education, 2006. Smaller Learning Communities Program. U.S. Department of Education, Washington, DC.
Watt, T.T., 2003. Are small schools and private schools better for adolescents' emotional adjustment? Sociology of Education 76, 344–367.