Can Small High Schools of Choice Improve Educational Prospects for Disadvantaged Students?

Howard S. Bloom
Rebecca Unterman

Abstract

This paper provides rigorous evidence (for 12,130 participants in a series of naturally occurring randomized lotteries) that a large-scale high school reform initiative (New York City's creation of 100+ small high schools of choice between 2002 and 2008) can markedly and consistently increase high school graduation rates (by 9.5 percentage points overall and for many different student subgroups) for a large population of educationally and economically disadvantaged students of color without increasing annual school operating costs. These findings are directly relevant to current debates by policymakers and practitioners about how to improve the educational prospects of disadvantaged students in the United States. © 2014 by the Association for Public Policy Analysis and Management.

INTRODUCTION

America's public high schools have been struggling for decades. For example, recent findings from the Common Core of Data indicate that in the 2008–09 school year a quarter of the nation's high school students failed to graduate after four years. Furthermore, this proportion has only decreased by 2.9 percentage points since the 2001–02 school year (Chapman, Ifill, & Ramani, 2011). Similarly, data from the Program for International Student Assessment indicate that in 2009, the reading skills of U.S. 15-year-olds ranked 14th in the world and had not improved statistically significantly during the previous decade. The situation is even worse for students of color. For example, in the 2008–09 school year, 36.5 percent of black high school students and 34.1 percent of Hispanic high school students failed to graduate after four years (Stillwell, Sable, & Plotts, 2011). With a modern labor market that requires increasingly sophisticated cognitive skills (Murnane, Willett, & Levy, 1995) and an increasingly competitive international economy (Wade, 2011), these findings are cause for serious concern by policymakers.

To date, three main approaches have been used to address this concern. One approach operates *outside* of existing school districts by offering school vouchers or creating charter schools that compete with districts for students and resources. While a lottery-based study of the District of Columbia Opportunity Scholarship Program (OSP) has shown that receiving a voucher has a positive effect on students' probability of graduating high school and on their reading achievement levels, there

Journal of Policy Analysis and Management, Vol. 33, No. 2, 290–319 (2014)

Published by Wiley Periodicals, Inc. View this article online at wileyonlinelibrary.com/journal/pam
DOI:10.1002/pam.21748

was no effect on math levels and more evidence is needed to determine whether the intervention is effective for all student subgroups and at scale (Wolf et al., 2013). Although there is some rigorous evidence about the effectiveness of charter schools, the strongest of which is based on randomized lotteries (Abdulkadiroglu et al., 2011; Dobbie & Fryer, 2009; Gleason et al., 2010; Hoxby & Murarka, 2009; McClure et al., 2005), the findings from this research are mixed. Furthermore, only 90 charter schools, or 2 percent of the national total, have been studied through lotteries (Betts & Atkinson, 2012; Witte et al., 2007); and only a small fraction of these 90 charters are high schools.

A second approach to high school improvement operates *within* school districts by comprehensively reforming their failing schools, in many cases by creating small learning communities within existing schools and rearranging other features of their instructional programs. Considerable research exists about the effectiveness of alternative high school reform models, and like the research on charter schools, it has produced mixed results. Reform models studied include among others, Career Academies (Kemple, 2008), Talent Development Schools (Kemple, Herlihy, & Smith, 2005), First Things First (Quint et al., 2005), and Project GRAD (Snipes, Holton, & Doolittle, 2006). Of these, only Career Academies was evaluated by a randomized trial.

A third approach to high school improvement that also works *within* school districts, and is the subject of the present paper, is the creation of new small high schools. In some cases, multiple new small schools are created to replace a large failing one that was closed, and the new schools are located together in the building of the closed school (Darling-Hammond, Ancess, & Ort, 2002; Kahne, Sporte, & Easton, 2005; Quint et al., 2010). In other cases, new small high schools are created in free-standing locations (e.g., Kahne, Sporte, & Easton, 2005 and Quint et al., 2010).

Following the work of early small school reformers, in the late 1990s, the Bill and Melinda Gates Foundation (BMGF) began making sizable investments to create small schools across the country. By 2005, the foundation had granted roughly $647 million to over 100 school districts. These grants varied substantially—from roughly $500,000 to the Grand Rapids school district to over $79 million (from 2000 through 2003) to 16 organizations in New York City that were selected to help start "small, academically challenging, but nonselective high schools" (Edweek, 2005). With these and other funds, initial efforts to create new small schools in New York City began in the mid-1990s and accelerated rapidly after 2002, as did similar efforts in cities, such as Cincinnati; Seattle; Washington, DC; Oakland; Milwaukee; and Chicago. Notably, Chicago created over 90 charter, contract, and district-run new small schools as part of its Renaissance 2010 Initiative (Wasley et al., 2000).

The present study uses naturally occurring randomized lotteries to estimate the effects on student outcomes of New York City's small high school initiative. The core research questions addressed by the study are (1) to what extent did New York City's new small high schools of choice increase graduation rates for their students beyond what these rates would have been had they attended another type of local public high school, (2) to what extent does the answer to this question vary across policy-relevant student subgroups, and (3) to what extent are high school costs for students in new small schools of choice higher than they would have been had they attended another type of local public high school? The remainder of this paper describes existing research on small schools, New York City's small public high schools of choice, the research design and analytic procedure used to estimate effects of these schools, and the findings that were obtained.

PRIOR SMALL-SCHOOL RESEARCH

The large and longstanding literature on small schools began with Barker and Gump's seminal book in 1964, *Big School, Small School: High School Size and Student Behavior*. Since then, much has been written about rationales for and against small schools. Among the more systematic reviews of this research are those by Leithwood and Jantzi (2009) and Cotton (1996). This research suggests that some small schools can produce beneficial effects on student achievement (Lee & Smith, 1997), promotion and attendance (Fowler & Walberg, 1991), dropout and graduation rates (Pittman & Haughwout, 1987; McMullen, Sipe, & Wolf, 1994), social relations (Lee, Bryk, & Smith, 1993), and school engagement (Klem & Connell, 2004). Small schools also have been shown to have more cohesive academic curricula and smaller classes (Monk, 1987). Lee and Smith (1997) argue that the "ideal high school . . . enrolls between 600 and 900 students" (p. 205). In contrast, using two different statistical approaches with data from the Educational Longitudinal Study of 2002, Schneider, Wyse, and Keesler (2007) find no effects of small schools on 12-graders' performance on standardized math tests, their expectations for post-secondary education, the number of colleges to which they apply or the type of college to which they apply.

The basis for most of these findings is a comparison of levels or changes in educational outcomes for students who attend small schools with those for students who do not, controlling statistically for observed differences between the groups. However, this approach cannot control for unobserved differences in factors such as parental engagement and student motivation that could influence future educational outcomes. Thus, differences in observed future outcomes might in part reflect preexisting differences in unobserved student characteristics.

To address this shortcoming, two recent papers used an instrumental variables strategy to estimate the effects of small high schools (Barrow, Claessens, & Schanzenbach, 2009; Schwartz, Stiefel, & Wiswall, 2012). The instruments used were based on the distance between students' homes and their nearest small high school. Findings from Barrow, Claessens, and Schanzenbach (2009) suggest that the 22 new small high schools created in Chicago since 2002 (with total enrollments of less than 600 students each) do not improve students' cognitive skills, but might improve their noncognitive skills. Findings from Schwartz, Steifel, and Wiswall (2012) suggest that small high schools created in New York City (defined as those with no more than 550 students) during the 1990s *reduce* graduation rates, while those created after 2002 *increase* graduation rates. Although the estimation strategy used by these two studies is probably stronger than that used for earlier research, it still is susceptible to bias from associations that might exist between the locations of students' homes and unobserved factors that influence their school outcomes. These unobserved factors could include, for example, characteristics of the neighborhoods where students live.

THE INTERVENTION

In 2002, New York City embarked on a large-scale series of educational reforms. With the granting of mayoral control, the New York City Department of Education (NYCDOE) was centralized and individual schools were granted greater autonomy over their curriculum, professional development, and teacher hiring, in exchange for stricter accountability standards and public performance assessments (O'Day, Bitter, & Gomez, 2011). Additional reforms included, among other things, a system of transfer schools for high school students who were close to dropping out, programs to produce new school leaders, and a "Fair Student Funding" model that provides resources to schools based in part on the educational needs of their students (NYCDOE, 2013a). At the heart of the high school component of these efforts were three interrelated reforms that were implemented between 2002 and 2008: (1) a district-wide high school choice process for all rising ninth graders; (2) closure of 31 large failing high schools (with an average graduation rate of 40 percent); and (3) creation of more than 200 new small high schools.

The new small schools were intended to serve students in some of the district's most disadvantaged communities and are located mainly in neighborhoods where large failing high schools had been closed. About 47 percent of these new schools are in the Bronx, 33 percent are in Brooklyn, and almost all are in low-income neighborhoods. In addition, most of the new small schools are located at or near large failing schools that were closed. In Appendix A, we present the location of small schools of choice (SSCs) that were opened and large, failing high schools that were closed between the fall of 2002 and the fall of 2008.[1]

The new schools are small, academically nonselective, and because they were created to provide a realistic choice for students with widely varying academic backgrounds, we call them "small schools of choice" or SSCs.[2] In addition to their uniformly small size (serving roughly 100 to 120 students per grade) and location mainly in disadvantaged neighborhoods, SSCs have the following defining traits.

First, SSCs were developed through a competitive proposal process that was designed to stimulate innovative ideas from a range of community stakeholders, such as parents, teachers, other school leaders, and local nonprofit organizations, and ensure that school founders met prespecified conditions regarding academic rigor and sustained personal relationships among students and faculty. For example, in 2008 SSC proposals were assessed with respect to 10 criteria, almost all of which are directly related to the SSC core principles of "rigor, relevance and relationships." The following are examples of these criteria (NYCDOE, 2008, p. 142).

- High expectations for all students and a standard-based, academically rigorous curriculum that connects what students learn with college and career goals,
- connections between what students learn in school to their lives and communities through internships, mentoring experiences, and service learning opportunities, and
- a structure that fosters the development of authentic, sustained, caring, and respectful relationships between teachers and students and among staff members.

Second, unlike in some cities, where small high schools were created by reconfiguring large, existing schools into smaller units in the same buildings with largely the same teachers and students, the typical SSC was created from scratch with a principal, teachers, and students who were new to the school. Furthermore, SSCs have organizational structures that are specifically designed to promote close and sustained interactions between teachers and students. For example, most SSCs were founded with community partners who offer students relevant learning opportunities inside and outside the classroom, and provided school faculty with additional staffing support and resources during start-up. In addition, SSCs are organized around small educational units that are designed to give students a better chance

[1] All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's Web site and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

[2] Other new, small secondary schools created by New York City during this period include 38 general high schools for grades 9 to 12, which screen students based on their academic backgrounds; 21 transfer schools that are designed to help students who are overage for grade and undercredited; 33 middle/high schools for grades 6 to 12 or 7 to 12; and one specialized high school for high-performing students.

of being known by their teachers and other adults in the building. For example, many SSCs have student advisory structures that carve out time at least once a week for teachers to meet with students and discuss youth development issues, such as maintaining self-esteem, building supportive peer relationships, and facing adversity. Furthermore, many SSCs organize teachers into grade-level teams and provide common planning time for teachers to discuss students' progress and solve problems for students who are struggling.

Third, SSCs received start-up funding from philanthropic organizations plus technical assistance from the NYCDOE and intermediary organizations, such as New Visions for Public Schools, the Urban Assembly, the Institute for Student Achievement, and the College Board. This assistance helped to facilitate school leadership development, staff hiring, and program startup. For example, most SSCs began with only a ninth-grade class, adding a subsequent grade each year. In addition, in the early years of this district-wide reform, SSCs received special allowances with respect to serving English Language Learners and special education students during their first two years of start-up. Furthermore, the creation of SSCs was supported by a consortium of funders led by the BMGF, the Carnegie Corporation of New York, and the Open Society Institute. Also, SSCs were implemented in collaboration with the United Federation of Teachers and the Council of School Supervisors and Administrators.[3] Lastly, SSCs are part of the NYCDOE and SSC teachers are members of New York City's teacher union.

OUR RESEARCH DESIGN

As noted, the present paper examines the extent to which enrolling in an SSC instead of another public high school operated by the NYCDOE increases rates of high school graduation for the study sample overall and for a wide range of student subgroups within it. These findings are based on data from naturally occurring lotteries produced as a by-product of New York City's High School Application Processing System (HSAPS), which was implemented in spring 2004. HSAPS currently assigns more than 90 percent of New York City's 90,000 entering ninth graders each year to over 400 public high schools.[4] Through HSAPS, students submit up to 12 high school choices in rank order, schools submit their criteria for prioritizing students, and students are assigned to schools based on this information. Three quarters of students currently receive one of their top three choices (Quint et al., 2010).

How HSAPS Creates SSC Lotteries

SSC priorities are based *solely* on students' geographic proximity and whether they are known to the SSC. Most SSCs have two categories of geographic proximity based on whether or not students live in the New York City borough where the SSC is located.[5] Within each of these geographic categories, SSCs give priority to students who are known to them over students who are not known to them. Students

[3] As described by Randi Weingarten at a forum on small schools sponsored by the Alliance for Excellent Education in November of 2010, the United Federation of Teachers participated in the planning and approval process for many of New York City's new small schools.

[4] HSAPS has three rounds of student placement. The second round places the vast majority of New York City's incoming ninth graders and is the basis for the present analysis. The first round assigns students who applied to one of New York City's eight specialized competitive high schools. The third round assigns students who are not placed by the first two rounds or who dispute their second round placement (see Abdulkadiroglu, Pathak, & Roth, 2005, for further details).

[5] Some SSCs have three geographic priorities: (1) residents of a nearby catchment area, (2) other residents of their borough, and (3) other residents of New York City.

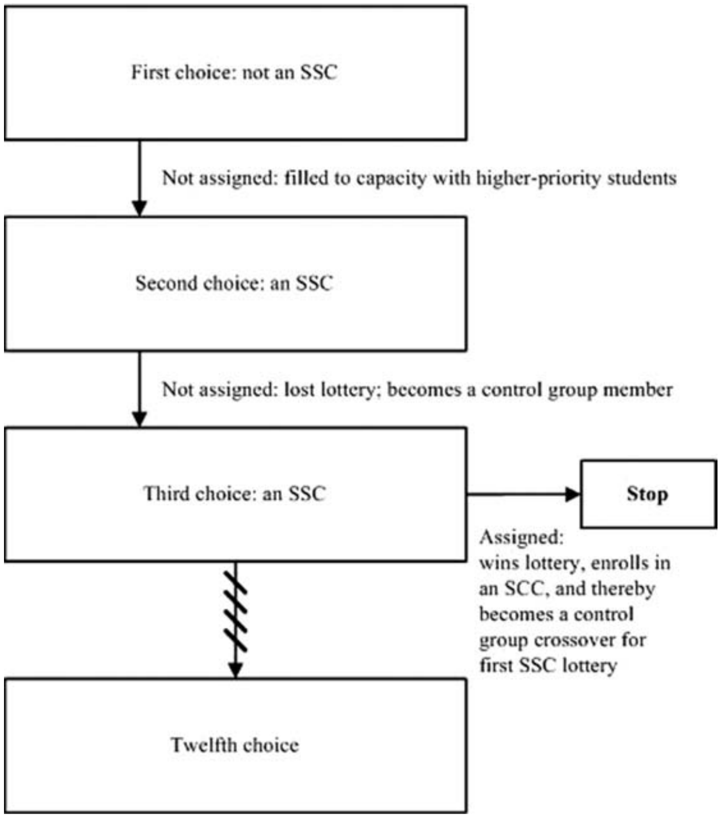


Figure 1. HSAPS Assignment Process for a Hypothetical Student.

can become known to an SSC in many ways, such as by contacting it, visiting it, or meeting with its representative at a high school fair.

When an SSC priority cell is oversubscribed by students from an annual cohort of incoming ninth graders, HSAPS *randomly* determines who is and is not assigned to it, thereby creating the statistical equivalent of a lottery for that cell (and thus for the SSC). We refer to these as "SSC lotteries." Consider the following intuitive description of this process. It begins with the HSAPS algorithm *randomly* determining the order in which it will assign students to high schools. The first students assigned are the most likely to receive their first choice school because no high schools are filled to capacity when HSAPS assigns them.[6] As high schools start to fill up, however, their student priorities begin to take effect. For example, if the first-choice high school of a student who is currently being placed by HSAPS is filled by students with equal or higher priority, the current student is not assigned to that school. Instead he is assigned to his next most-preferred school with available space. This process continues until HSAPS assigns all incoming ninth graders in its queue.

Figure 1 illustrates this process for a hypothetical student. In the example, HSAPS does not assign the student to his first-choice school (which is not an SSC) because he was chosen after the school was filled by other students with the same or a

6 This is the case as long as students are not subsequently bumped from the school by students who are assigned later by HSAPS and have higher priority for that school.

higher priority for that school (other types of schools have other priority systems). The student also is not assigned to his second-choice school (an SSC) either because the school was already filled by students with a higher priority for it or because it was already filled by students with the same priority for it, but an earlier place in the random order by which HSAPS assigns students to schools.[7] If the school was already filled by students with the same priority, the current student would have been in the equivalent of a lottery for that SSC (described below) and lost it. If the SSC were filled by students with a higher priority, the current student would not have been in a lottery. Instead he would have been preempted by students with higher priority for that school. Assume that the student lost a lottery for this SSC and thus became a member of its control group for the present study.

The student is finally assigned by HSAPS to his third-choice school, which is an SSC. If he were assigned to this school because it was *not* oversubscribed by students with the same or a higher priority for it, he would *not* have been in a lottery for it. If he was assigned to the school and it was oversubscribed by students with the same priority for it, he would have been in a lottery for this school. Assume that the student was in a lottery for this SSC and won it: He lost a lottery for an SSC that was his second-choice school, won a lottery for an SSC that was his third-choice school, and was assigned to the latter SSC. Consequently, the student was a control group crossover for his first SSC lottery (which was for his second-choice school).

To demonstrate how HSAPS creates a lottery for an SSC when it is oversubscribed, Figure 2 illustrates HSAPS assignment for a hypothetical SSC that can accommodate 120 incoming ninth graders. HSAPS will attempt to assign to this SSC all students who list it as one of their choices and are *not* assigned to a more-preferred choice. For example, students who list the SSC as their third choice and do not receive their first or second choice (like the hypothetical student in the example) are considered for assignment to the SSC. In contrast, students who list the SSC as their third choice, but receive their first or second choice are not considered for assignment to the SSC.

The hypothetical SSC has 360 students who list it as a choice and do not receive a more-preferred choice. Hence the SSC has 360 potential assignees. Eighty of them are from Priority Cell 1 because they live in the school's borough and are known to it, 160 are from Priority Cell 2 because they live in the school's borough and are not known to it, 50 are from Priority Cell 3 because they do not live in the school's borough and are known to it, and 70 are from Priority Cell 4 because they do not live in the school's borough and are not known to it. Given the school's capacity of 120 entering ninth graders and the fact that it accepts students according to their priority cell, it can accept all 80 students from Priority Cell 1 plus the first 40 students randomly assigned by HSAPS from Priority Cell 2. It cannot accept the last 120 students randomly assigned from Priority Cell 2 or any students from Priority Cells 3 and 4.

Because Priority Cell 2 is oversubscribed, the 160 students in this cell are effectively lottery participants. The first 40 of these participants randomly selected for school assignment by HSAPS (through their random order in the HSAPS student assignment queue) win the present lottery and are assigned to the SSC. The last 120 participants randomly selected by HSAPS for school assignment lose the present lottery and are not assigned to the SSC. Within this cell, only the random order in which HSAPS selects students for school assignment determines who wins or loses

[7] If the SSC were filled by students with a *lower* priority, HSAPS would assign the current student to the SSC and bump the student with lowest priority that was most recently assigned to it. The student that was bumped would then be assigned to his next most preferred school with available space.

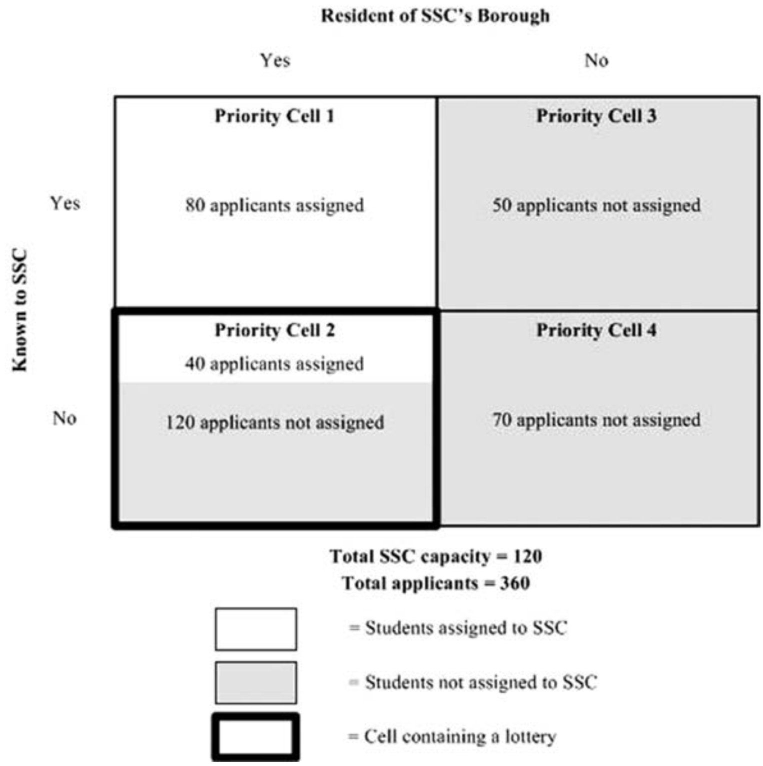


Figure 2. HSAPS Assignment Process for a Hypothetical SSC.

the lottery. Thus at this point in the process, *students' rank-ordered preferences for the SSC do not influence whether they are assigned to it.* For example, it does not matter whether one lottery participant listed the SSC as his first choice and another listed it as his 12th choice. The only thing that determines which participants win the lottery and which ones lose it is the random order in which HSAPS assigns students (i.e., who gets there first).

Note that no lottery exists for Priority Cell 1 of this SSC because all of its potential assignees are assigned to the SSC. Furthermore, no lottery exists for Priority Cells 3 and 4 because all of their potential assignees are preempted by students with a higher priority for that SSC. Because only one cell has a lottery for a given SSC in a given year, the winners and losers of each lottery are matched or blocked by their priority for that SSC. In this way each SSC lottery produces a naturally occurring randomized trial for each SSC that is oversubscribed in a given year.

Before proceeding further it is important to note that some participants in the current lottery in Figure 2 might have lost an HSAPS lottery for a more preferred school. This school might have been an SSC or some other type of high school that had an oversubscribed admissions stratum that year (limited unscreened schools, zoned schools, unscreened schools, or multiprogram high schools with both screened and unscreened programs in the same building).[8] For example, the hypothetical student

[8] Go to http://schools.nyc.gov/ChoicesEnrollment/High/Admissions/default.htm for a discussion of the admissions process for these schools. We include instances of oversubscription in SSCs plus all of these other types of schools as HSAPS lotteries.

described in Figure 1 lost a lottery for his second-choice school, which was an SSC, and then won a lottery for his third-choice school, which was an SSC. Hence, this student had participated in a prior HSAPS lottery. Because outcomes of prior lotteries are determined before outcomes of a current lottery, whether or not students were in a prior HSAPS lottery is an exogenous baseline characteristic of current lottery participants. This characteristic can be used to produce two exogenously defined subgroups of current lottery participants—those that were in a prior HSAPS lottery and those that were not. As described below, this fact plays an important role in the construction of the present sample.

How SSC Lotteries Produced the Present Sample

The present sample is comprised of SSC lottery participants from three annual cohorts of first-time ninth graders who entered high school in the fall of 2005, 2006, or 2007 and met six criteria. First, the present sample only includes a participant in a given SSC lottery if it was *his first HSAPS lottery*.[9] For example, recall that HSAPS did not assign the hypothetical student in Figure 1 to his first-choice school (which was not an SSC) because it was already filled to capacity by students who had a higher priority for that school. Hence, the student was not in an HSAPS lottery for his first-choice school. The first HSAPS lottery for this student was for his second-choice school (an SSC), and the present sample would include him in that SSC lottery. Because the student lost this lottery, he is a control group member for it in the present analysis. This student would not be included in the lottery sample for his third-choice school (an SSC) because it was his second HSAPS lottery. Because the student won this second lottery and enrolled in its SSC, he is a control group crossover for his first lottery. Note that for any given current HSAPS lottery, whether or not a participant had been in a logically prior lottery is determined before the outcome of the current lottery. Hence, this student characteristic provides an exogenous way to determine subgroups of participants in the current lottery. Therefore, omitting participants in a current lottery who had participated in a logically prior lottery is an exogenous exclusion that does not affect the randomness of assignment of the remaining lottery participants.

The following additional exogenous exclusions were applied to produce the present sample. (1) Eighth graders from parochial or private schools were excluded because their HSAPS identifiers do not provide a reliable way to track them through high school enrollment.[10] (2) Eighth graders who were missing scores for both of their New York state eighth-grade tests (reading and math) were excluded because most were from outside of New York State and their HSAPS identifiers do not provide a reliable way to track them through high school enrollment. (3) After the previous exclusions were applied, 30 lotteries with 1,374 students no longer had any treatment-group members or no longer had any control group members.[11] These

[9] Appendix B of Bloom, Thompson, and Unterman (2010) describes a potential theoretical problem that could arise from including participants in an SSC lottery for whom this is not their first HSAPS lottery. This problem could arise if a student participated in a logically prior HSAPS lottery for an SSC or some other type of high school. Bloom and Unterman (2013) demonstrate that this potential problem does not exist in practice because estimates of SSC effects obtained by including these lottery participants are virtually identical to those obtained by not including them.

[10] When private and parochial students participate in HSAPS, they are often given a unique identifier that exists solely for the high school application process. This identifier can differ from that received if the student subsequently enrolls in a NYCDOE high school. Thus, the two datasets cannot be merged and used for this analysis.

[11] Each of these lotteries had only one or two treatment-group members or one or two control group members *before* the sample exclusions were invoked.

incomplete lotteries could not be used to study the effects of SSCs. (4) There were 26 lotteries with 1,864 students for SSCs that opened in a later year or in a location, which differed from that which was planned.[12] Many, if not all, of these students were thus not assigned to high school based on these lotteries. (5) Six SSC lotteries with 110 students had zero compliance with their SSC assignment. This means that the percentage of lottery winners who enrolled in an SSC equaled the percentage of control group members who did so.[13] These lotteries thus provide no information about the effect of enrolling in an SSC.

The preceding exclusions produced a final lottery sample of 14,969 students from 199 lotteries for 84 SSCs. From this sample, 2,553 students were lost to the present analysis because they were missing four-year follow-up data, and 286 students from seven lotteries were missing because their lottery became incomplete when students with missing follow-up data were omitted.[14] The resulting four-year follow-up sample contains 12,130 students from 192 lotteries for 84 SSCs, which represents an overall student attrition rate of 19.0 percent.

An estimate (which accounts for the blocking of randomization by lottery) of the treatment- and control-group difference in attrition rates indicates that this difference is only 0.53 percentage points and is not statistically significant (P-value = 0.481).[15] Thus, winning or losing an SSC lottery does not affect the rate of student attrition. Furthermore, Table 1 presents estimates (which also account for the blocking of randomization by lottery) that demonstrate that mean baseline characteristics are virtually identical for SSC lottery winners and control group members in the four-year follow-up sample (Table B.1 in Appendix B demonstrates that this is also the case for the final lottery sample[16]). Hence, student attrition does not appear to have affected the internal validity of the present analysis.

Data Sources

The primary sources of data for the present analysis are information for individual students from the HSAPS and other NYCDOE administrative records. In addition, publically available data on school characteristics were obtained from New York State School Report Cards and the U.S. Department of Education's Common Core of Data.

HSAPS data were used to identify students who participated in an SSC lottery, to determine the school to which they were assigned and the school in which they enrolled, to describe students' characteristics, and to compare these characteristics

[12] These schools did not open the following fall with their HSAPS Round Two school code or HSAPS Round Two assignment students. In these instances, the school opening was delayed a year or the school location was moved and the school admitted students using the HSAPS Round Three assignment process.
[13] These very small lotteries had total samples ranging from four to 24 students and control groups ranging from one to three students. All winners and control group members from these lotteries enrolled in an SSC.
[14] Six of these lotteries had only one to three treatment-group members or one to three control group members *before* sample exclusions were invoked. One of these lotteries had 42 control group members and 91 treatment-group members.
[15] The treatment- and control-group difference of attrition rates or difference of means for a baseline characteristic was estimated as the value of β_X in the regression, $X_i = \sum_{j=1}^{J} \alpha_j \cdot I_{ji} + \beta_X \cdot T_i + v_i$, where X_i is a zero/one attrition indicator or a specific baseline characteristic for student i, the I_{ji} are a series of zero/one lottery indicators for the j lotteries in the study sample, T_i is a zero/one indicator of winning or losing one's first HSAPS lottery, and v_i is a random error.
[16] All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's Web site and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

Table 1. Baseline characteristics of first-time SSC lottery participants: four-year follow-up sample, cohorts 1 to 3.

Characteristic (%)	SSC lottery winners	Control group members	Estimated difference	*P*-value for estimated difference
Race/ethnicity				
Hispanic	43.7	45.3	−1.6	0.148
Black	46.0	44.9	1.1	0.314
American Indian	0.4	0.5	0.0	0.811
White	5.0	5.1	−0.1	0.809
Asian	3.6	3.1	0.5	0.257
Male	47.1	45.6	1.5	0.182
Eligible for free/reduced-price lunch	83.4	83.2	0.3	0.796
English language learner	6.1	5.9	0.2	0.762
Special education	5.4	5.5	−0.1	0.889
Overage for eighth grade	16.7	17.1	−0.4	0.686
Eighth-grade reading proficiency				
Did not meet standards (level 1)	6.6	5.9	0.7	0.261
Partially met standards (level 2)	61.8	63.2	−1.4	0.208
Fully met standards (level 3)	30.5	30.0	0.5	0.642
Met standards with distinction (level 4)	1.1	0.8	0.3	0.229
Eighth-grade math proficiency				
Did not meet standards (level 1)	19.1	18.5	0.6	0.524
Partially met standards (level 2)	47.5	47.2	0.2	0.834
Fully met standards (level 3)	31.6	32.4	−0.8	0.451
Met standards with distinction (level 4)	1.9	1.9	−0.1	0.865
Missing eighth-grade pretests				
Math proficiency	0.7	0.8	−0.1	0.543
Reading proficiency	3.0	2.9	0.1	0.815
Sample size (total = 12,130)	5,020	7,110		
Number of lotteries (total = 192)				
Number of SSCs (total = 84)				

Sources: Findings are based on data from the NYCDOE HSAPS, NYCDOE state test data for eighth graders in 2004–05 (cohort 1), 2005–06 (cohort 2), and 2006–07 (cohort 3) plus NYCDOE enrollment files for the 2005–06 to 2010–11 school years.

A chi-square test was used to assess the statistical significance of the overall difference between lottery winners and control group members reflected by the full set of baseline characteristics in the table. The resulting chi-square value is not statistically significant (*P*-value = 0.430).

for lottery winners and control group members. These data include student's rank-ordered preference lists of high schools, their baseline characteristics, and SSCs' priority categories for them.

Students' school records data were used to construct follow-up measures of academic progress, which are the basis for estimates of SSC effects. This information includes enrollment and course credits earned plus results of state Regents examinations. Middle-school data on eighth-grade state test scores in reading and math were obtained for baseline comparisons of SSC lottery winners and control group members.

Data on direct service expenditures per pupil, the most comprehensive measure of high school costs that is available for cross-school comparisons, were obtained for each of the schools in the present analysis from annual NYCDOE School-Based Expenditure Reports.

Analytic Approach

The present analytic approach focuses on the effect of enrolling in an SSC, not the effect of winning one's first SSC lottery. This is because the effect of winning one's first SSC lottery does not have a useful causal interpretation. Specifically, if a student wins his first SSC lottery, he is assigned by HSAPS to that SSC. However, if he loses that lottery it is still possible for him to be assigned by HSAPS to another SSC that is lower on his school preference list. Thus, the difference between winning and losing one's first SSC lottery is not the same as the difference between being assigned by HSAPS to an SSC and not being assigned by HSAPS to an SSC. Consequently, the effect of winning one's first SSC lottery is not the same as the effect of being assigned by HSAPS to an SSC.

We estimate the causal effect of enrolling in an SSC using an instrumental variables approach that is widely employed for randomized trials to estimate effects of receiving a treatment. The approach uses random assignment to treatment or control status in one's first SSC lottery as an instrumental variable or instrument, to predict SSC enrollment, which in turn is used to estimate the causal effect of enrolling in an SSC.[17]

This approach produces an estimate of the average effect of enrolling in an SSC for *students who did so because they won their first SSC lottery*. This type of causal effect is typically referred to as a local average treatment effect and the individuals to which this affect applies are often referred to as compliers (because they comply with their assigned treatment).[18] However, compliance is an ambiguous concept for students who lose one SSC lottery, win another, and consequently enroll in an SSC. Such students are noncompliers for their first lottery and either compliers or noncompliers for their second or later lottery or lotteries. To avoid this ambiguity, we use the term *target SSC enrollee* to designate students for whom SSC enrollment effects are estimated. We chose this term because these students are the target of estimation. It should be noted that target SSC enrollees are sample members who are compliers with respect to their first SSC lottery.

In a multiblock randomized trial, like that used for the present analysis (with SSC lotteries as blocks), it is often useful to specify a separate instrument for each block. These instruments can be created by interacting a zero/one indicator for treatment assignment (T_i) with a zero/one indicator for each block (I_{ji}). The present instruments are valid because they are randomized and strong because they are highly predictive of SSC enrollment. To increase precision, students' scores on their eighth-grade New York State tests of math and reading were included as baseline covariates.[19] Two-stage least squares was then used to estimate a model like that described by equations 1 and 2 below.

[17] For discussions of this approach to analyzing multisite randomized trials see Gennetian et al. (2005); Ludwig and Kling (2007); and Kling, Liebman, and Katz (2007); for analyzing randomized studies in general, see Angrist, Imbens, and Rubin (1996) and Angrist and Pischke (2009); for analyzing lottery-based studies in particular, see Abdulkadiroglu et al. (2011).

[18] See Angrist, Imbens, and Rubin (1996).

[19] Approximately 0.83 percent of sample members were missing their eighth-grade math pretest score and 2.44 percent were missing their eighth-grade reading pretest scores. These missing values were imputed from a single model-based replicate of a multiple imputation model using all available baseline and follow-up data. SAS PROC MI was used for this purpose.

First Stage: SSC Enrollment as a Function of First SSC Lottery Assignment

$$E_i = \sum_{j=1}^{J} \pi_j \cdot I_{ji} + \sum_{j=1}^{J} \gamma_j \cdot T_i \cdot I_{ji} + \theta_M \cdot S_{Mi} + \theta_R \cdot S_{Ri} + w_i \tag{1}$$

where

$E_i = 1$ if student i enrolled in an SSC at any time during the four-year follow-up period and 0 otherwise,[20]
$I_{ji} =$ if SSC lottery j was the first HSAPS lottery for student i and 0 otherwise,
$T_i = 1$ if student i won his SSC lottery and 0 otherwise,
S_{Mi} and $S_{Ri} =$ student i's eighth-grade scores on New York State tests of math and reading,
$w_i =$ a random error that is distributed independently and identically across students within lotteries.

Second Stage: High School Outcome as a Function of Predicted SSC Enrollment

$$Y_i = \sum_{j=1}^{J} \alpha_j \cdot I_{ji} + \delta \cdot \hat{E}_i + \phi_M \cdot S_{Mi} + \phi_R \cdot S_{Ri} + e_i \tag{2}$$

where

$Y_i =$ the outcome for student i,
$\hat{E}_i =$ the predicted value of SSC enrollment for student i from the estimated first-stage equation,
$e_i =$ a random error that is clustered by the first school that students entered after their lottery.

The estimated value of δ is a consistent estimate of the average effect of enrolling in an SSC for target SSC enrollees.[21]

FINDINGS

This section presents key findings from the present analysis.

Four-Year Graduation Effects Overall and by Student Cohort

Findings in Table 2 indicate that on average, enrolling in an SSC increased four-year graduation rates by 9.5 percentage points (to 70.4 percent for target SSC enrollees from 60.9 percent for their control group counterparts).[22] This estimated effect

[20] Although we define enrolling in an SSC as enrolling in *any* SSC, 87.6 percent of the first SSC lottery winners who enrolled in an SSC enrolled in the SSC *for that lottery*. Thus, enrolling in an SSC almost always means enrolling in the SSC for the lottery that a student won.
[21] Estimates of second-stage standard errors were adjusted to account for uncertainty in predicted SSC enrollment $\hat{E}_i$.
[22] Target SSC enrollees and their control group counterparts are by definition, students who comply with their first SSC lottery assignment and who also would have complied if their lottery assignment had been different. Appendix C describes how outcome levels for target SSC enrollees were estimated. The estimated level for control group counterparts equals the difference between the estimated mean level for target SSC enrollees and the estimated effect of enrolling in an SSC. All appendices are available at

Table 2. Estimated effects of SSCs on four-year high school graduation rates by student cohort: cohorts 1 to 3.

Cohort	Target SSC enrollees (%)	Control group counterparts (%)	Estimated effect (%)	*P*-value for estimated effect	Sample size
Cohort 1 (2004 to 2005)	66.6	58.3	8.3*	0.014	4,473
Cohort 2 (2005 to 2006)	70.4	59.2	11.2**	0.000	3,995
Cohort 3 (2006 to 2007)	74.6	65.1	9.5**	0.001	3,662
Cohorts 1 to 3	70.4	60.9	9.5**	0.000	12,130

Sources: Findings are based on data from the NYCDOE HSAPS for eighth graders in 2004–05 (cohort 1), 2005–06 (cohort 2), and 2006–07 (cohort 3), plus data from NYCDOE files on student attendance, course credits, Regents examination scores, administrative transactions, and enrollment for the 2005–06 to 2010–11 school years.

Note: A two-tailed *t*-test was used to assess the statistical significance of each estimated SSC effect with significance levels indicated as ** = 1 percent; * = 5 percent. Variation in estimated SSC effects across the three cohorts was not statistically significant (*P*-value for chi-square test = 0.798).

Table 3. Estimated effects of SSC enrollment on four-year graduation rates accounting for student attrition: cohorts 1 to 3.

Graduated from high school (%)	Target SSC enrollees	Estimated effect	Control group counterparts	*P*-value for estimated effect	Sample size
Four-year follow-up sample					
No imputed outcomes	70.4	60.9	9.5**	0.000	12,130
Final SSC lottery sample					
Model-based imputed outcomes	64.6	56.1	8.5**	0.000	14,969
Imputed outcomes set to zero	59.3	50.7	8.5**	0.000	14,969

Sources: Findings are based on data from the NYCDOE HSAPS for eighth graders in 2004–05 (cohort 1), 2005–06 (cohort 2), and 2006–07 (cohort 3), plus data from NYCDOE files on student attendance, course credits, Regents examination scores, administrative transactions, and enrollment for the 2005–06 to 2010–11 school years.

Note: A two-tailed *t*-test was applied to the estimated effect. Statistical significance levels are indicated as ** = 1 percent; * = 5 percent.

varies across cohorts from 8.3 to 11.2 percentage points.[23] Thus, SSC effects on four-year high school graduation rates were successfully replicated for three annual student cohorts. This result is especially striking given that graduation rates had been rising district-wide during the study period and thus at the schools against which SSCs were compared. Kemple (2013) provides evidence of this rising tide, as does the fact that graduation rates for control group counterparts in Table 2 increased from 58.3 percent for the first cohort to 65.1 percent for the third cohort.

Findings in Table 3 demonstrate that student attrition is unlikely to markedly influence estimates of SSC effects on high school graduation rates, although accounting for attrition reduces estimates of the underlying graduation rates for target SSC enrollees and their control group counterparts. The first row in the

the end of this article as it appears in JPAM online. Go to the publisher's Web site and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

[23] This variation is not statistically significant at the 0.05 level (*P*-value = 0.809).

table repeats estimates of the average SSC effect on four-year graduation rates for the four-year follow-up sample of 12,130 students. The second and third rows report corresponding findings for the final lottery sample of 14,969 students, using two different approaches to impute missing follow-up data.

The first and most fully informed imputation approach is based on a single replicate of a multiple imputation model that uses all existing baseline and follow-up data for members of the final lottery sample plus their treatment- or control-group status to impute missing data. These findings indicate that enrolling in an SSC increased high school graduation rates by 8.5 percentage points (to 64.6 percent for target SSC enrollees from 56.1 percent for their control group counterparts). The estimated 8.5 percentage-point effect that accounts for sample attrition is very close to the estimated 9.5 percentage-point effect that does not account for sample attrition. However, the underlying graduation rates for target SSC enrollees and their control group counterparts are appreciably lower when sample attrition is taken into account than when it is not taken into account. This reflects the fact that students with missing four-year outcome data entered high school with much weaker graduation prospects than was the case for other students.[24]

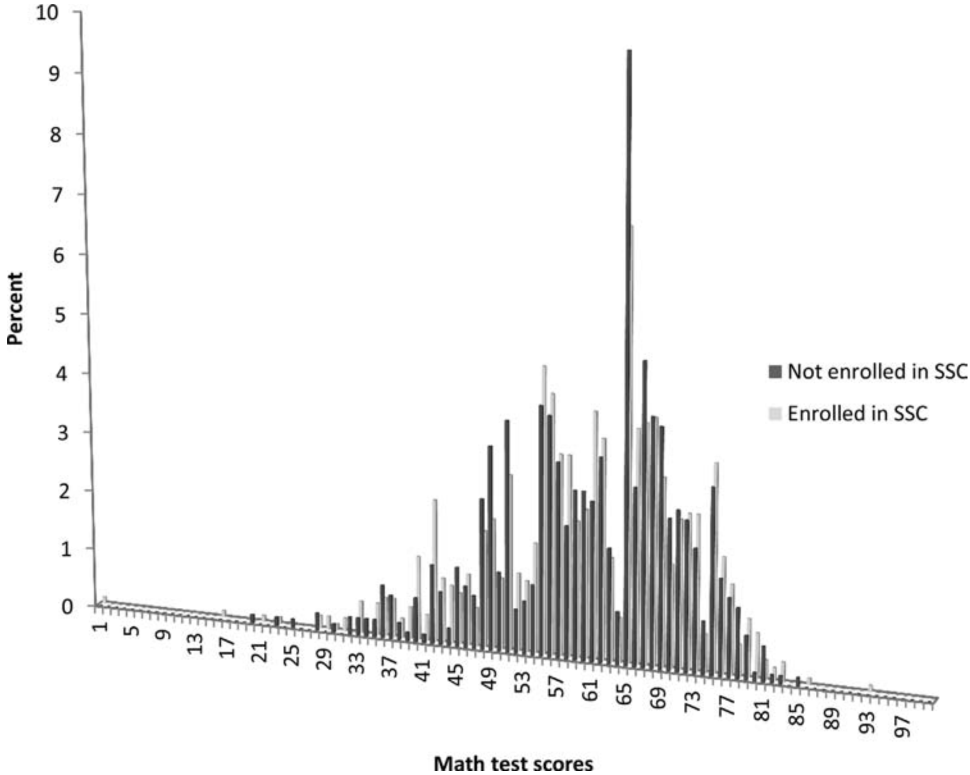
The third row in the table presents findings for the final lottery sample that assume no students with missing follow-up data graduated from high school in four years. This conservative assumption produces even lower estimates of graduation rates for target SSC enrollees and their control group counterparts. However, the estimated SSC effect is the same as that from model-based imputation, 8.5 percentage points.[25] Furthermore, both of these estimates are similar to that obtained without imputing missing follow-up data—9.5 percentage points.

A second test that was conducted to assess the robustness of the present findings was to reestimate the average effect of enrolling in an SSC on four-year graduation rates (for the four-year follow-up sample) using a single instrument (whether or not each student won his first SSC lottery) instead of using multiple instruments (one for each lottery). The resulting estimate was a 10.6 percentage point increase in the four-year graduation rate, which was highly statistically significant (P-value = 0.000). Hence, the central finding of the present paper is essentially the same regardless of whether it is estimated using multiple randomized instruments (as was done for all other findings in this paper) or a single randomized instrument.

Figure 3 presents a third robustness test. As described later, receiving a New York State Regents diploma, which is the standard high school credential in New York state, requires that, among other things, a student score at least 65 points on each of five Regents examinations: English language arts, mathematics, science, global history, and American history. Because teachers grade the Regents performance of students from their own school, there have been longstanding concerns about undue teacher influence on these scores. These concerns were heightened by recent research that documented a pronounced bulge in the number of scores that were at

[24] Specifically, 31.3 percent of students who were missing four-year outcome data entered high school overage for grade versus 18.7 percent for students in the four-year follow-up sample. After one year of high school, 28.9 percent of students who were missing four-year follow-up data (but had follow-up data for one year) were on track to graduate in four years versus 54.9 percent of students in the four-year follow-up sample.

[25] It is not plausible to assume that the graduation rate for attriters who were lottery winners is *lower* than that for attriters who were control group members because the best available evidence indicates that SSCs *increased* early progress toward graduation on average for the 94 percent of attriters with available data on this early outcome. Specifically, among four-year attriters with first year follow-up data, we estimate that 33.0 percent of target SSC enrollees were on track toward graduation versus 30.4 percent of their control group counterparts. This estimated advantage was not statistically significant (P-value = 0.435).



Sources: Findings are based on data from the NYCDOE HSAPS for eighth graders in 2004–05 (cohort 1), 2005–06 (cohort 2), and 2006–07 (cohort 3), plus data from NYCDOE files on student attendance, course credits, Regents examination scores, administrative transactions, and enrollment for the 2005–06 to 2010–11 school years.

Figure 3. Distribution of Scores on the New York State Math Regents Examination for SSC Enrollees and Non-SSC Enrollees in the Four-Year Follow-Up Sample.

or just above the threshold of 65 points (Dee et al., 2011; Martinez & McGinty, 2011). Figure 3 indicates that this bulge in math Regents scores is no greater for SSCs than for other New York City public high schools, and comparable findings are presented in Appendix D for the other four Regents examinations.[26] Thus, undue teacher influence on students' Regents examination scores is unlikely to be responsible for the positive SSC effect on graduation rates.

Four-Year Effects on Graduation by Diploma-Type and College Readiness

During the present study period, students in New York state could receive one of three types of high school diplomas: a local diploma (which had the least stringent requirements), a New York State Regents diploma (which is generally considered

[26] All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's Web site and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

Table 4. Estimated effects of SSCs on four-year high school graduation and college readiness: cohorts 1 to 3.

Outcome (%)	Target SSC enrollees	Control group counterparts	Estimated effect	*P*-value for estimated effect
Graduation				
Graduated from high school	70.4	60.9	9.5**	0.000
Local diploma granted	17.0	14.7	2.3	0.145
Regents diploma granted	45.9	39.9	6.0**	0.007
Advanced Regents diploma granted	7.5	6.3	1.2	0.469
College readiness				
English Regents examination score of 75 or above	40.2	33.4	6.8**	0.002
Math A Regents examination score of 75 or above	24.6	24.7	0.0	0.989

Sources: Findings are based on data from the NYCDOE HSAPS for eighth graders in 2004–05 (cohort 1), 2005–06 (cohort 2), and 2006–07 (cohort 3), plus data from NYCDOE files on student attendance, course credits, Regents examination scores, administrative transactions, and enrollment for the 2005–06 to 2010–11 school years.

Note: A two-tailed *t*-test was used to assess the statistical significance of each SSC estimated effect with significance levels indicated as ** = 1 percent; * = 5 percent.

to be the standard graduation credential), and a New York State Advanced Regents diploma (which has the most stringent requirements). Table 4 indicates that most of the increase in graduation rates caused by SSCs is due to a 6.0 percentage-point increase in receipt of New York State Regents diplomas.[27] To obtain this type of diploma, students must score at least 65 points on each of the five required New York State Regents examinations (as noted above) and pass all courses required by the state. A much smaller portion of the increase in graduation rates caused by SSCs is due to a 2.3 percentage-point increase in receipt of local diplomas, which were phased out for future student cohorts. An even smaller portion of the increase in graduation rates caused by SSCs is due to a 1.2 percentage-point increase in receipt of Advanced Regents diplomas that are received by very few students who apply to SSCs.

Table 4 also indicates that enrolling in an SSC increased students' college readiness in English as measured by the percentage that scored at least 75 points on the English Regents examination. This threshold is used by the City University of New York (CUNY) to exempt incoming students from taking remedial English. Enrolling

[27] Table E.1 in Appendix E reports estimates of the average effect of winning a student's first SSC lottery on high school graduation and college readiness. All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's Web site and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

Table 5. Estimated effects of SSCs on five-year graduation rates by student cohort and diploma type: cohorts 1 and 2.

By cohort or diploma type (%)	Target SSC enrollees	Control group counterparts	Estimated effect	*P*-value for estimated effect	Sample size
By cohort					
Cohort 1 (2004–05)	71.4	64.3	7.1*	0.027	4,500
Cohort 2 (2005–06)	76.1	65.9	10.3**	0.000	4,021
Cohorts 1 and 2	73.6	64.7	8.9**	0.000	8,521
By diploma type for cohorts 1 and 2					
Local diploma	23.1	21.0	2.2	0.320	8,521
Regents diploma	44.1	38.0	6.2*	0.011	8,521
Advanced Regents diploma	6.3	5.6	0.7	0.654	8,521

Sources: Findings are based on data from the NYCDOE HSAPS for eighth graders in 2004–05 (cohort 1) and 2005–06 (cohort 2) plus data from NYCDOE files on student attendance, course credits, Regents examination scores, administrative transactions, and enrollment for the 2005–06 to 2010–11 school years.

Note: A two-tailed t-test was used to assess the statistical significance of each estimated SSC effect with significance levels indicated as ** = 1 percent and * = 5 percent. The difference between estimated SSC effects for cohorts 1 and 2 is not statistically significant (P-value for t-test = 0.463).

in an SSC increased this rate by 6.8 percentage points (to 40.2 percent for target SSC enrollees from 33.4 percent for their control group counterparts). In contrast, enrolling in an SSC had no effect on college readiness in mathematics as measured by the percentage of students that scored at least 75 points on their Regents mathematics examination.[28] In addition, target SSC enrollees were much more likely to be college ready in English than in mathematics (40.2 percent vs. 24.6 percent).

Five-Year Graduation Effects by Student Cohort and Diploma Type

Table 5 demonstrates that SSC effects on high school graduation and college readiness are sustained after five years of follow-up. Findings in the top panel indicate that on average, SSCs increased five-year graduation rates for students in the first two cohorts by 8.9 percentage points (to 73.6 percent for target SSC enrollees from 64.7 percent for their control group counterparts).[29] Estimates of this effect for each cohort (7.1 and 10.3 percentage points) are about 1 percentage point smaller than their four-year counterparts. This reflects the fact that during the fifth follow-up year, graduation rates for target SSC enrollees rose by between 5 and 6 percentage points, while those for their control group counterparts rose by between 6 and 7 percentage points.

Findings in the bottom panel of Table 5 indicate that the effect of SSCs on five-year graduation rates is due primarily to an increase in receipt of Regents diplomas as was the case for their four-year effect. Thus the key result after five years, as

[28] In 2012, CUNY raised its threshold for exempting students from taking remedial mathematics to 80 points. The estimated SSC effect on the percentage of students who surmounted this threshold is 0.0 percentage points.
[29] Five-year follow-up data are not yet available for cohort 3.

graduation rates for sample members continued to rise, is roughly the same as that after four years.

Four-Year Graduation Effects by Student Subgroup

Findings in Table 6 for subgroups within the study's first three student cohorts indicate that SSCs increased four-year graduation rates for many different types of students.[30] The first findings in the table are for subgroups defined by students' prior academic proficiency, as measured by their eighth-grade performance on New York state tests of reading and math. Proficiency levels 1 and 2 are considered by the state to be below grade level, whereas proficiency levels 3 and 4 are considered to be at or above grade level. The findings indicate that SSCs markedly increased four-year graduation rates for students in the first three proficiency levels of both subjects, with estimated effects ranging from 9.8 to 11.8 percentage points. This is especially striking given the enormous variation across these subgroups in their counterfactual graduation rates for control group counterparts (from 30.2 to 74.7 percent for reading and from 36.9 to 76.8 percent for math). SSCs had no effect on graduation rates for the few sample members who are in the top prior proficiency levels because almost all of them would graduate without attending an SSC.

The next findings in the table indicate that SSCs increased graduation rates for subgroups of students that varied in terms of their eligibility for free/reduced price lunch and their gender and race/ethnicity. These findings are followed by those for subgroups defined by students' eligibility for special education or English language learner services.[31] As can be seen, SSCs appear to increase graduation rates for special education students by 13.8 percentage points, although this estimate is not statistically significant given the subgroup's limited sample size. SSCs appear to increase graduation rates for English language learners by 4.9 percentage points, although given the small size of this subgroup and its modest estimated effect, this finding is not statistically significant.

Table 6 also indicates that SSC effects for students who were sufficiently motivated and informed to make themselves known to their SSC during the high school choice process were not larger than those for students who did not make themselves known. In addition, the table indicates that students who felt strongly enough to make a given SSC their first choice school did not experience a larger effect of attending an SSC than did students who ranked SSCs lower on their school preference list.

The SSC Treatment Contrast

On average, SSC enrollees attended an SSC for 3.34 school years during the first four years after they entered high school. Table 7 compares key features of these SSCs with those of schools attended by control group counterparts. These differences represent dimensions of the SSC *treatment contrast* that were measured by estimating SSC enrollment effects on each school characteristic using the

[30] To estimate SSC effects for a student subgroup, the sample for each SSC lottery was stratified by the student characteristic that defined the subgroup. Because not all lotteries include students from each stratum, different subgroup samples represent different combinations of lotteries.

[31] SSC lotteries include special education students who can be taught in a regular classroom setting. Special education students classified by the NYCDOE as requiring collaborative team teaching services or self-contained classes are not part of SSC lotteries because they are not assigned to schools in round two of HSAPS.

Table 6. Estimated effects of SSCs on four-year graduation rates for student subgroups: cohorts 1 to 3.

Student characteristic (%)	Target SSC enrollees	Control group counter-parts	Estimated effect	*P*-value for estimated effect	Sample size
Eighth-grade reading proficiency[†]					
Did not meet standards (level 1)	40.3	30.2	10.1	0.152	817
Partially met standards (level 2)	68.1	58.3	9.8**	0.000	6,452
Fully met standards (level 3)	85.9	74.7	11.2**	0.000	4,348
Met standards with distinction (level 4)	88.5	90.0	−1.5	0.698	513
Eighth-grade math proficiency[†]					
Did not meet standards (level 1)	47.2	36.9	10.2*	0.043	3,578
Partially met standards (level 2)	68.1	56.3	11.8**	0.000	5,707
Fully met standards (level 3)	88.4	76.8	11.5	0.000	2,548
Met standards with distinction (level 4)	98.8	100[a]	−1.2	0.790	297
Low-income status					
Eligible for free/reduced-price lunch	68.6	57.4	11.2**	0.000	7,418
Not eligible for free/reduced-price lunch	73.2	66.2	7.0*	0.011	4,712
Race/ethnicity, by gender					
Black male	65.5	52.0	13.5**	0.002	2,300
Black female	72.8	64.7	8.0*	0.042	2,917
Hispanic male	64.5	57.2	7.3	0.068	2,745
Hispanic female	73.2	62.9	10.3**	0.002	2,930
Other male	83.2	77.6	5.6	0.197	552
Other female	87.6	78.1	9.5	0.154	510
Known or unknown to SSC					
Known	72.9	64.9	8.0**	0.002	6,823
Unknown	63.6	51.6	12.1**	0.000	5,307
Choice level (of 12) at which enrollee participated in lottery					
First choice	71.8	63.3	8.5**	0.001	5,688
Second choice	66.9	60.0	6.8*	0.045	2,689
All other choices	70.2	55.1	15.2**	0.000	3,753
Special education status					
Eligible for services	62.8	48.9	13.8	0.074	725
Not eligible for services	70.7	61.5	9.3**	0.000	11,405
English language learner					
Eligible for services	63.9	59.0	4.9	0.418	843
Not eligible for services	70.8	61.3	9.5**	0.000	11,287

Sources: Findings are based on data from the NYCDOE HSAPS for eighth graders in 2004–05 (cohort 1), 2005–06 (cohort 2), and 2006–07 (cohort 3), plus data from NYCDOE files on student attendance, course credits, Regents examination scores, administrative transactions, and enrollment for the 2005–06 to 2010–11 school years.

Note: A two-tailed *t*-test was used to assess the statistical significance of each estimated SSC effect with significance levels indicated as ** = 1 percent; * = 5 percent. A chi-square test was used to assess the statistical significance of variation in estimated SSC effects across subgroups within a given dimension, with significance levels indicated as: † = 5 percent.

Table 7. School characteristics for target SSC enrollees and control group counterparts: four-year follow-up sample, cohorts 1 to 3.

School characteristic	Target SSC enrollees	Control group counterparts	Estimated difference	*P*-value for estimated difference
School age (%)				
School opened since 2002	100.0	7.9	92.1	0.000
School was reformed/ restructured since 2002	0.0	14.1	−14.1	0.000
School established before 2002	0.0	67.5	−67.5	0.000
School size				
Number of first-time Ninth graders enrolled in ninth grade	109.8	467.1	−357.4	0.000
Small—550 students or less (%)	100.0	22.1	77.9	0.000
Medium—551 to 1,400 students (%)	0.0	21.8	−21.8	0.000
Large—more than 1,400 students (%)	0.0	56.1	−56.1	0.000
Medium and large with Small learning communities (%)	0.0	11.7	−11.7	0.000
Characteristics of the school's first-time ninth graders (%)				
Race/ethnicity				
Hispanic	43.8	48.3	−4.5	0.000
Black	46.7	41.9	4.8	0.000
White	5.4	5.0	0.4	0.164
Asian	3.6	4.3	−0.6	0.011
Eligible for free/reduced-price lunch	80.9	82.5	−1.5	0.017
Special education	11.8	13.8	−2.0	0.000
English language learners	6.3	11.9	−5.5	0.000
Scored at or above eighth-grade level in reading	30.3	25.1	5.2	0.000
Scored at or above eighth-grade level in math	32.7	27.4	5.4	0.000
Overage for eighth grade	20.2	25.0	−4.8	0.000
Teacher characteristics (%)				
Less than 3 years of teaching experience	37.8	22.8	15.0	0.000
Doctorate or master's degree plus 30 hours	21.7	30.9	−9.2	0.000
Total number of student observations = 12,130				

Sources: Findings are based on data from the NYCDOE HSAPS for eighth graders in 2004–05 (cohort 1), 2005–06 (cohort 2), and 2006–07 (cohort 3) plus data from the New York State Report Card for the 2002–03 to 2010–11 school years, NYCDOE enrollment and course credit files for the 2005–06 to 2010–11 school years, and NYCDOE school-level administrative records for the 2002–03 to 2010–11 school years.

Note: A two-tailed *t*-test was applied to the estimated difference. Statistical significance levels are indicated as ** = 1 percent; * = 5 percent.

instrumental variables model that produced our estimates of SSC effects (equations 1 and 2).

Note first that, as expected, target SSC enrollees attended schools that are much newer and smaller than those attended by their control group counterparts. Note next that target SSC enrollees attended schools with first-time ninth-grade peers that are demographically and economically similar to those of control group counterparts. The proportions of peers that were special education students were also similar. However, the first-time ninth-grade peers of target SSC enrollees appear to be about 5 percentage points less likely to be English language learners, less likely to have scored below grade level on state eighth-grade tests of reading and math, and less likely to have been overage for grade in eighth grade. Thus, SSCs may provide a modest advantage in terms of student peers.

Lastly, Table 7 indicates that teachers of target SSC enrollees have fewer years of experience and are less likely to have graduate school credits or degrees. This is consistent with existing perceptions that SSCs attract teachers who are new to the teaching profession.[32]

The Study Sample in Context

Table 8 compares background characteristics of target SSC enrollees with those of all first-time ninth graders in the study's SSCs who were assigned there by HSAPS during the sample intake period[33] and with those of all first-time ninth graders in New York City public high schools during this period. These findings indicate that target SSC enrollees are quite similar to all HSAPS enrollees in the study SSCs with respect to their race/ethnicity, gender, eligibility for free or reduced-price lunch, English language learner status, likelihood of being overage for grade, and prior performance on eighth-grade state tests. They only differ with respect to the percentage of students who live in the Bronx and the percentage of students with special education status (the latter of which reflects HSAPS reporting limitations that make it possible to only identify—and thus include in the study—special education students taught in mainstream classrooms).

Now consider how target SSC enrollees compare to all entering ninth graders in New York City public schools. First, note that a greater percentage of target SSC enrollees are black or Hispanic and eligible for free or reduced-price lunches than is the case for all entering ninth graders district-wide. Thus, SSCs are serving a disproportionately large number of students of color who live in poverty. Next, note that the percentage of target SSC enrollees that are overage for grade upon entering high school and their scores on eighth-grade state tests of reading and math are similar to those for all entering ninth graders in New York City. However, both in the study sample and citywide, well over half of incoming ninth graders are performing below grade level when they enter high school.

[32] These comparisons are made in terms of school-level teacher characteristics for each sample member, because it is not possible to obtain information on his or her specific teachers.
[33] Because not all SSCs had an HSAPS lottery for all three student cohorts in the present analysis and because even when an SSC had a lottery, not all of the student assigned to it by HSAPS were in this lottery, the number of students assigned by HSAPS to the SSCs in the present study is far larger than the number of lottery participants in the present sample.

Table 8. Baseline characteristics of target SSC enrollees, all HSAPS enrollees in study SSCs and all first-time ninth graders in New York City: cohorts 1 to 3.

Characteristic (%)	Target SSC enrollees	All HSAPS enrollees in study SSCs	All first-time ninth-grade students
Race/ethnicity			
Hispanic	47.9	48.3	37.7
Black	42.7	45.6	34.3
Other	8.3	6.1	28.0
Male	48.0	51.3	51.4
Eligible for free/reduced-price lunch	83.7	84.0	75.9
Special education	6.0	15.4	13.9
English language learner	7.0	8.3	11.2
Overage for eighth grade	21.3	25.1	22.8
Eighth-grade reading proficiency			
Did not meet standards (level 1)	7.2	10.6	10.5
Partially met standards (level 2)	63.2	63.2	51.2
Fully met standards (level 3)	29.2	25.5	35.1
Met standards with distinction (level 4)	0.5	0.7	3.2
Eighth-grade math proficiency			
Did not meet standards (level 1)	18.1	22.6	18.4
Partially met standards (level 2)	45.1	44.5	36.2
Fully met standards (level 3)	34.3	30.9	36.5
Met standards with distinction (level 4)	2.5	2.0	8.9
Borough (home residence)			
Bronx	54.9	49.1	23.3
Brooklyn	28.6	31.2	32.4
Manhattan	8.2	11.2	12.2
Queens	6.0	7.3	26.8
Staten Island	2.4	1.2	5.3
Total number of student observations = 12,130			

Sources: Findings are based on data from the NYCDOE HSAPS, NYCDOE state test data for eighth graders in 2004–05 (cohort 1), 2005–06 (cohort 2), and 2006–07 (cohort 3) plus data from NYCDOE enrollment files from the 2005–06 to 2010–11 school years.

Cost Considerations[34]

To help assess the policy relevance of the educational benefits produced by SSCs relative to the schools against which they were compared, Table 9 reports their relative costs. Discussions in the literature about potential economies of scale in the production of education suggest that small schools are more expensive to operate than large schools (e.g., Andrews, Duncombe, & Yinger, 2002). On the other hand, there is some evidence that when small schools are themed they might be no more expensive than larger comprehensive schools (Stiefel et al., 2009).

For the present analysis, high school costs are presented in terms of two measures of *direct service expenditures per pupil*, which is a comprehensive measure of

[34] Findings in this section were produced in collaboration with Professor Robert Bifulco from the Maxwell School of Citizenship and Public Affairs at Syracuse University. Further details and a more extensive analysis will be presented in a forthcoming paper by Bifulco, Bloom, and Unterman (2014).

Table 9. Effects of SSC enrollment on five-year direct service expenditures per entering ninth grader: five-year follow-up sample, cohorts 1 to 2.

	Target SSC enrollees	Control group counterparts	Estimated effect	Estimated effect (%)	*P*-value for estimated effect
Direct service expenditures					
Year one	11,934	11,955	−20.29	−0.2	0.943
Year two	12,708	12,962	−253.93	−2.0	0.337
Year three	13,926	14,398	−471.79	−3.3	0.056
Year four	14,887	15,431	−543.81	−3.5	0.052
Year five	3,740	5,029	−1289.44	−25.6**	0.000
Total	57,195	59,774	−2579.25	−4.3*	0.024
Adjusted direct service expenditures					
Year one	12,506	12,181	325.47	2.7	0.256
Year two	13,274	13,159	115.37	0.9	0.654
Year three	14,493	14,579	−86.56	−0.6	0.726
Year four	15,387	15,583	−196.04	−1.3	0.494
Year five	3,796	5,069	−1273.07	−25.1**	0.001
Total	59,457	60,571	−1114.83	−1.8	0.328
Sample size	3544	4977			

Sources: Findings are based on data from the NYCDOE HSAPS for eighth graders in 2004–05 (cohort 1) and 2005–06 (cohort 2) plus data from the NYCDOE School Based Expenditures Reports for school years 2005–06 through 2010–11.

Notes: A two-tailed t-test was applied to the estimated effect. Statistical significance levels are indicated as ** = 1 percent; * = 5 percent.

school- and district-level expenditures for services provided directly to students during the school year and school day. The first measure, direct service expenditures per pupil, is based on information for each NYCDOE high school obtained from the annual NYCDOE School Based Expenditure Report for school years 2005–06 through 2010–11. To place these findings in perspective note that per pupil direct service expenditures for the average school in our cost analysis (SSCs and other high schools) during the 2005–06 through 2010–11 school years were $12,650, $13,485, $14,693, $15,595, $16,239, and $15,853, respectively.

The second measure, *adjusted direct services per pupil*, modifies the first measure to account for the fact that SSC teachers are less experienced on average and thus, according to the city's collective bargaining agreement, are paid lower salaries on average than are teachers in other NYCDOE high schools. This second measure values the time of all teachers in all high schools according to the district-wide average teacher salary. This was done to remove the influence of teacher experience levels on school costs (see Appendix F[35]).

Results in the table were obtained by (1) assigning a per pupil expenditure to each entering ninth-grade sample member for each of five follow-up years based on the high school he was enrolled in during October of each school year (the official New York state reporting date for enrollment), and (2) estimating the effect of SSC enrollment on this expenditure using the statistical model that was used to estimate

[35] All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's Web site and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

SSC effects on students' academic outcomes. These findings represent estimated direct service expenditures *per entering ninth grader* in our first two student cohorts (five-year follow-up data are not yet available for the third cohort).

Results for the two cost measures tell the same story: Direct service expenditures are no higher (and perhaps are slightly lower) for target SSC enrollees than for their control group counterparts. During each of the first four years of high school there was no systematic difference between these expenditures for the two groups. This probably reflects the fact that 22.1 percent of control group counterparts attended a small school, 21.8 percent attended a medium-size school, and 11.7 percent attended a large school with small learning communities (see Table 7). Thus, a large fraction of control group counterparts attended schools that were relatively small or had small instructional units.

During the fifth year of high school, average direct service expenditures per entering ninth grader in the study sample dropped precipitously for both target SSC enrollees and their control-group counterparts because only a small fraction of them enrolled for a fifth year. Interestingly, fifth-year expenditures per entering ninth grader were about one third *higher* for control group counterparts than for SSC enrollees ($5,029 vs. $3,740). This is because fifth-year enrollment rates were about one-third higher for control group counterparts (26.4 percent vs. 19.7 percent).[36] And this in turn, probably reflects that fact that roughly one third more control group counterparts were available for a fifth year of high school because they did not graduate in their fourth year (see Table 2).

Table 9 is missing two potentially important SSC costs: (1) the one-time cost of creating an SSC and (2) the ongoing costs of human and financial resources contributed to SSCs by their external partners. As noted earlier, SSCs received special supports during their first few years of operation. For example, throughout this period the BMGF aimed to donate roughly $100,000 a year to SSC planning teams for their school planning year and the first four years of their school's operation. The exact amount that SSCs received from BMGF is difficult to document because these funds were given to intermediary organizations to support their work and also passed through to SSCs. Thus, although internal BMGF documents suggest that roughly 135 million dollars went to support small schools initiatives in New York City between the fall of 2000 and the fall of 2007, over a third of these funds were for purposes that complemented but were not directly related to school start-up or early operation. In addition, other foundations invested in this movement and the NYCDOE itself gave special funds to new schools created during this period. Thus, a full accounting of these funds from multiple sources is beyond the scope of the present paper and may not even be possible. However, it is important to note that while these start-up resources might have been substantial, they represent a one-time cost for each new SSC that should be amortized over many years of its subsequent operation. Thus, it is unclear how much difference an accounting of these costs would make to the findings that we present.

The best available information about the ongoing costs of contributions to SSCs from their external partners is a budgetary survey of 13 SSC external partners conducted by New Visions for Public Schools. The survey results indicate that on average, during the 2004–05 school year, these external partners contributed financial and in-kind resources that were worth roughly $275,000 per SSC (Soler-McIntosh, Carrion, & Guntan, 2007). For a typical SSC with roughly 440 students, this equals $625 per student, which is only about 5 percent of annual direct service expenditures. Thus, adding these off-the-book costs for SSCs does not change our conclusion

[36] These rates were estimated using our standard impact model (equations 1 and 2).

about their overall operating costs. Furthermore, since other NYCDOE schools can have external partners, only considering external contributions for SSCs will overstate the difference in this cost component between SSCs and other NYCDOE high schools. This makes it even more likely that if a full accounting of these costs were possible, it would not change our conclusion.

The negligible additional costs of SSCs are especially notable given the dearth of rigorous evidence on other inexpensive effective academic attainment and dropout prevention programs for high school students. For example, the What Works Clearinghouse has only given a *Meets Standards* rating to two such interventions to date—Check and Connect and *ALAS*. While each had positive early effects on students' transition into and progress through high school, they cost $1,400 per student per year and $1,185 per participant per year (U.S. Department of Education, 2006a, 2006b).

CONCLUSIONS, CAVEATS, AND FURTHER QUESTIONS

The preceding findings provide rigorous evidence that in a relatively short period of time (roughly seven years), with sufficient organization and resources, an existing school district can implement a complex high school reform that markedly improves graduation rates for a large population of low-income, disadvantaged students of color. In this way, the New York City SSC experience provides an existence proof that successful large-scale high school reform is possible.

As noted, estimates of SSC effects are based on follow-up data for a very large sample (12,130 students) and for an especially rigorous research design that is based on a series of naturally occurring randomized lotteries. Hence, these findings provide unusually valid and reliable evidence about the effectiveness of a major high school intervention. Furthermore, instead of relying on a single test score to measure high school effectiveness, the present findings reflect SSC effects on a combination of students' course attainment and success on multiple New York State Regents examinations that are required for a high school diploma. In addition, these estimates are quite large in magnitude. For example, the estimated 9.5 percentage point increase in four-year graduation rates produced by SSCs is equivalent to roughly one half of the gap in graduation rates between white students and students of color in New York City.[37] Furthermore, this effect is relative to schools that exist currently, which do not include the 31 large failing high schools that were closed, and was sustained as graduation rates were rising over time in the 200+ high schools against which SSCs were compared.

Of further note is the fact that SSCs markedly increased graduation rates for students who entered high school with widely varying graduation prospects. For example, when students were placed into subgroups according to their performance on eighth-grade New York state reading tests, graduation rates for control group counterparts ranged from 30.2 percent for students in the lowest reading level to 74.7 percent for students in the third highest level.[38] Nonetheless, SSCs increased graduation rates for students in all three categories by between 9.8 and 11.2 percentage points. Furthermore, the findings indicate that SSCs increased graduation rates for students who were and were not eligible for free and reduced-price lunches; students who were male and female; black and Hispanic; students who did and did not make themselves known in advance to their SSC; students for whom their SSC

[37] On average, across cohorts that entered high school in 2005–06 through 2007–08, the graduation gap between white students and students of color was approximately 19 percentage points (NYCDOE, 2013b).
[38] Very few control group counterparts were in the highest level and 90.2 percent of them graduated.

was their first-choice school, their second-choice school, or their third- through 12th-choice school; students who were eligible for special education services; and students who were eligible for English language learner services (although the last two estimates are not statistically significant).

The magnitude, consistency, and robustness of the estimated SSC effects are especially impressive given the unusually large scale of the SSC initiative and the highly disadvantaged population of students that it serves. For example, the 84 SSCs studied serve about 37,000 students, which is equivalent in size to the entire high school population of the Dallas Independent Secondary School District, the 14th largest secondary school district in the United States. In addition, New York City's SSC initiative to create district-run new small schools is far larger than small school initiatives in any other U.S. city. Furthermore, present estimates of SSC effects represent findings for a student population that is 90 percent black or Hispanic, 84 percent eligible for free or reduced-price lunch, 17 percent overage for grade in eighth grade, 68 percent performing below eighth-grade level in reading, and 66 percent performing below eighth-grade level in math. Lastly, the additional benefits produced by SSCs do not appear to have been at the expense of higher annual operating costs.

However, when considering these findings one should also consider the extent to which they generalize to other settings. As noted, New York City's SSCs were developed from scratch through a demanding proposal process. This differs markedly from many small schools created elsewhere by reconfiguring existing schools into smaller units with many of the same teachers and students. In addition, New York City's SSCs received technical assistance from the NYCDOE and intermediary organizations that were often experienced at launching new schools. Finally, in contrast to charter schools or some other small school initiatives, SSCs were implemented in collaboration with the United Federation of Teachers and the Council of School Supervisors and Administrators. Other school districts looking to implement SSCs should assess whether they have the same human and political capital to draw upon.

Lastly, it is important to note that the present findings raise as many questions as they answer. Perhaps the single most important question is the following: What are the active ingredients that enable SSCs to increase academic attainment and achievement for their students? Because SSCs differ from other high schools in so many ways, it remains to be seen which of these differences are most instrumental to their success.

HOWARD S. BLOOM is Chief Social Scientist, MDRC, 16 E. 34th Street, 19th Floor, New York, NY 10016. (e-mail: Howard.Bloom@mdrc.org).

REBECCA UNTERMAN is a Research Associate, MDRC, 16 E. 34th Street, 19th Floor, New York, NY 10016. (e-mail: Rebecca.Unterman@mdrc.org).

ACKNOWLEDGMENTS

Research for this paper was funded by the Bill & Melinda Gates Foundation. The authors thank David Silver and Fannie Tseng, our project officers, plus Adam Tucker, Sara Allan, and Steve Cantrell from the foundation for their support of this research. We also thank staff members from the New York City Department of Education who are too numerous to name for making this research possible. In addition, we thank Richard Murnane, Sean Reardon, Brian Jacob, Gordon Berlin, James Kemple, Saskia Levy Thompson, and Pei Zhu for their invaluable feedback plus Micah DeLaurentis, Rebecca Coven, Gary Reynolds, and Luis Lopez Vivas for helping to produce the present paper.

REFERENCES

Abdulkadiroglu, A., Pathak, P. A, & Roth, A. E. (2005). The New York City high school match. American Economic Review, 95, 364–367.

Abdulkadiroglu, A., Angrist, J. D., Dynarski, S. M., Kane, T. J., & Pathak, P. A. (2011). Accountability and flexibility in public schools: Evidence from Boston's charters and pilots. Quarterly Journal of Economics, 126, 699–748.

Andrews, M., Duncombe, W., & Yinger, J. (2002). Revisiting economies of size in American education: Are we any closer to a consensus? Economics of Education Reviews, 21, 245–262.

Angrist, J. D., & Pischke, J. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton, NJ: Princeton University Press.

Angrist, J. D., Imbens, G. W., & Rubin, D. (1996). Identification of casual effects using instrumental variables (with discussion). Journal of the American Statistical Association, 91, 444–472.

Barrow, L., Claessens, A., & Schanzenbach, D. W. (2009). The impact of small schools in Chicago: Assessing the effectiveness of Chicago's small high school initiative. Unpublished manuscript. Chicago, IL: University of Chicago.

Betts, J. R., & Atkinson, R. C. (2012). Better research needed on the impact of charter schools. Science 335, 171–172.

Bifulco, R., Bloom, H. S., & Unterman, R. (2014). Estimating the effects of New York City's small high schools of choice on students' high-school costs. Working paper. New York, NY: MDRC.

Bloom, H. S., & Unterman, R. (2013). Sustained progress: New findings about the effectiveness and operation of small public high schools of choice in New York City. New York, NY: MDRC.

Bloom, H. S., Thompson, S. L., & Unterman, R. with Herlihy, C., & Payne, C.F. (2010). Transforming the high school experience: How New York City's new small schools are boosting student achievement and graduation rates. New York, NY: MDRC.

Chapman, C., Ifill, J. L. N., & Ramani, A. K. (2011). Trends in high school dropout and completion rates in the United States: 1972–2009. Washington, DC: National Center for Education Statistics and the U.S. Department of Education.

Cotton, K. (1996). School size, school climate, and student performance: Close-up #20. Portland, OR: Northwest Regional Educational Laboratory.

Darling-Hammond, L., Ancess, J., & Ort, S. W. (2002). Reinventing high school: Outcomes of the coalition campus schools project. American Educational Research Journal, 39, 639–673.

Dee, T. S., Jacob, B. A., McCrary, J., & Rockoff, J. (2011). Rules and discretion in the evaluation of students and schools: The Case of the New York Regents examinations. Research Paper. New York, NY: Columbia Business School.

Dobbie, W., & Fryer R. G. Jr. (2009). Are high quality schools enough to close the achievement gap? Evidence from a social experiment in Harlem. NBER Working Paper No. 15473. Cambridge, MA: The National Bureau of Economic Research.

EdWeek. (2005, June 16). Major Gates Foundation grants to support small high schools. Education Week. Retrieved September 4, 2013, from http://www.edweek.org/ew/articles/2004/06/16/40chart.h23.html.

Fowler W. J. Jr., & Walberg, H. J. (1991). School size, characteristics, and outcomes. Educational Evaluation and Policy Analysis, 13, 189–202.

Gennetian, L. A., Morris, P., Bos, J., & Bloom, H. (2005). Constructing instrumental variables from experimental data to explore how treatments produce effects. In H. Bloom (Ed.), Learning more from social experiments: Evolving analytic approaches. New York, NY: Russell Sage Foundation, 75–114.

Gleason, P., Clark, M., Tuttle, C. C., & Dwoyer, E. (2010). The evaluation of charter school impacts: Final report. NCEE 2010-4029. Washington, DC: National Center for Education

Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Hoxby, C. M., & Murarka, S. (2009). Charter schools in New York City: Who enrolls and how they affect their students' achievement. NBER Working Paper No. 14852. Cambridge, MA: The National Bureau of Economic Research.

Kahne, J. E., Sporte, S. E., & Easton, J. Q. (2005). Creating small schools in Chicago: An early look at implementation and impact. Improving Schools, 8, 17–22.

Kemple, J. J. (2008). Career academies: Long-term impacts on labor market outcomes, educational attainment, and transitions to adulthood. New York, NY: MDRC.

Kemple, J. J. (2013). The condition of New York City high schools: Examining trends and looking toward the future. New York, NY: The Research Alliance for New York City Schools.

Kemple, J. J., Herlihy, C. M., & Smith, T. J. (2005). Making progress toward graduation. New York, NY: MDRC.

Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. Journal of School Health, 74, 262–273.

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. Econometrica, 75, 83–119.

Lee, V. E., & Smith, J. B. (1997). High school size: Which works best and for whom? Educational Evaluation and Policy Analysis, 19, 205–227.

Lee, V. E., Bryk, A. S., & Smith, J. B. (1993). The organization of effective secondary schools. Review of Research in Education 19, 171–267.

Leithwood, K., & Jantzi, D. (2009). A review of empirical evidence about school size effects: A Policy Perspective. Review of Educational Research 79, 464–490.

Ludwig, J., & Kling, J. R. (2007). Is crime contagious? Journal of Law and Economics, 50, 491–518.

Martinez, B., & McGinty, T. (2011, February 2). Students' Regents test scores bulge at 65. Retrieved November 22, 2013, from http://online.wsj.com/news/articles/SB10001424052748703445904576117793343465096.

McClure, L., Strick, B., Jacob-Almeida, R., & Reicher, C. (2005). The Preuss School at UCSD: School characteristics and students' achievement. San Diego, CA: The Center for Research on Educational Equity, Assessment and Teaching Excellence.

McMullen, B. J., Sipe, C. L., & Wolf, W. C. (1994). Charter and student achievement: Early evidence from school restructuring in Philadelphia. Philadelphia, PA: Center for Assessment and Policy Development.

Monk, D. H. (1987). Secondary school size and curriculum comprehensiveness. Economics of Education, 6, 137–150.

Murnane, R. J., Willett, J. B., & Levy, F. (1995). The growing importance of cognitive skills in wage determination. Review of Economics and Statistics, 77, 251–266.

New York City Department of Education (NYCDOE). (2008). New secondary school application, 2008 edition. New York, NY: Office of Portfolio Development.

New York City Department of Education (NYCDOE). (2013a). Fair student funding & school budget resource guide: FY 2013. New York, NY: Author.

New York City Department of Education (NYCDOE). (2013b). Graduation results. Retrieved November 22, 2013, from http://schools.nyc.gov.

O'Day, J. A., Bitter, C. S., & Gomez, L. M. (Eds.). (2011). Education reform in New York City: Ambitious change in the nation's most complex school system. Cambridge, MA: Harvard Education Press.

Pittman, R. B., & Haughwout, P. (1987). Influence of high school size on dropout rate. Education Evaluation and Policy Analysis, 9, 337–343.

Quint, J., Bloom, H. S., Black, A. R., & Stephens, L. with Akey, T. M. (2005). The challenge of scaling up educational reform: Findings and lessons from first things first. New York, NY: MDRC.

Quint, J., Smith, J., Unterman, R., & Moedano, A. (2010). New York City's changing high school landscape: High schools and their characteristics, 2002–2008. New York, NY: MDRC.

Schneider, B., Wyse, A. E., & Keesler, V. (2007). Is small really better? Testing some assumptions about high school size. In T. Loveless & F. M. Hess (Eds.), Brookings papers on education policy 2006/2007 (pp. 15–47). Washington, DC: Brookings Institution Press.

Schwartz, A. E., Stiefel, L., & Wiswall, M. (2012). Do small schools improve performance in large, urban districts? Causal evidence from New York City. Retrieved September 4, 2013, from http://ssrn.com/abstract=1986092 or http://dx.doi.org/10.2139/ssrn.1986092.

Snipes, J. C., Holton, G. I., & Doolittle, F. (2006). Charting a path to graduation: The effect of project GRAD on elementary school student outcomes in four urban school districts. New York, NY: MDRC.

Soler-McIntosh, J, Carrion, A., & Guntan, B. (2007). Partnership sustainability study. New York, NY: New Visions for Public Schools.

Stiefel, L., Schwartz, A. E., Iatarola, P., & Chellman, C. (2009). Mission matters: The cost of small high schools revisited. Economics of Education Review, 28, 585–599.

Stillwell, R., Sable, J., & Plotts, C. (2011). Public school graduates and dropouts from the Common Core Data: School year 2008–2009. Retrieved November 22, 2013, from http://nces.ed.gov/Pubsearch/Pubsinfo.Asp?Pubid=2011312.

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2006a). Dropout prevention: ALAS. Retrieved November 20, 2013, from http://ies.ed.gov/ncee/wwc/pdf/intervention_reports/WWC_Project_ALAS_100506.pdf.

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2006b). Dropout prevention: Check & connect. Retrieved November 20, 2013, from http://ies.ed.gov/ncee/wwc/pdf/intervention_reports/WWC_Check_Connect_092106.pdf.

Wade, R. H. (2011). Globalization, growth, poverty, inequality, resentment and imperialism. In J. Ravenhill (Ed.), Global political economy (pp. 373–409). Oxford, UK: Oxford University Press.

Wasley, P.A., Fine, M., Gladden, M., Holland, N. E., King, S. P., Mosak, E., & Powell, L. C. (2000). Small schools: Great strides. New York, NY: Bank Street College of Education.

Witte, J., Weimer, D., Shober, A., & Schlomer, P. (2007). The performance of charter schools in Wisconsin. Journal of Policy Analysis and Management, 26, 557–573.

Wolf, P. J., Kisida, B., Gutmann, B., Puma, M., Eissa, N., & Rizzo, L. (2013). School vouchers and student outcomes: Experimental evidence from Washington, DC. Journal of Policy Analysis and Management, 32, 246–270.

APPENDIX A

Location of SSCs That Were Opened and Large, Failing High Schools That Were Closed between the Fall of 2002 and the Fall of 2008

Figure A.1 reports the location of SSCs that were opened and large, failing high schools that were closed during the study period. In the map, each ring represents a new small school. When shaded rings surround black dots, this symbolizes an instance when a large high school was closed and an SSC was opened at the same address (on the same campus) or in close proximity. In the map, the shaded gray areas represent places where the median annual household income is less than $40,000. The observable overlap between these areas and the location of SSCs indicates that the majority of SSCs are located in low-income areas.

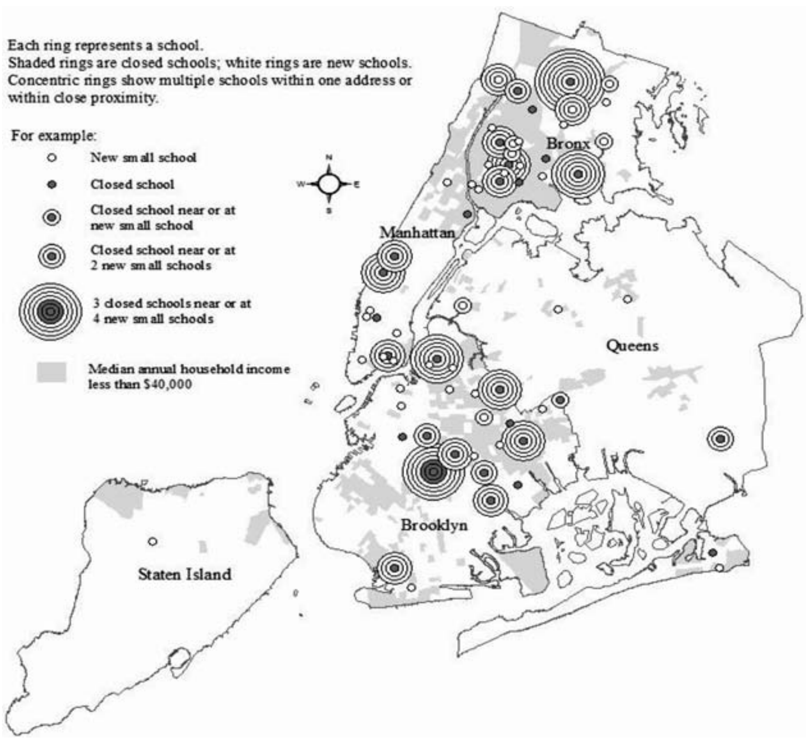


Figure A.1. Location of SSCs That Were Opened and Large, Failing High Schools That Were Closed between the Fall of 2002 and the Fall of 2008.

APPENDIX B

Comparing Baseline Characteristics of SSC Lottery Winners and Control Group Members

Table 1 in the paper compares baseline characteristics of SSC lottery winners (treatment-group members) and lottery losers (control group members) from the study's four-year follow-up sample. Table B.1 in this appendix reports the same

Table B.1. Baseline characteristics of first-time SSC lottery participants: final lottery sample, cohorts 1 to 3.

Characteristic (%)	SSC lottery winners	Control group members	Estimated difference	*P*-value for estimated difference
Race/ethnicity				
Hispanic	44.7	46.7	−2.0	0.074
Black	45.1	43.6	1.6	0.114
American Indian	0.5	0.5	0.0	0.862
White	4.7	4.5	0.2	0.682
Asian	3.5	3.4	0.1	0.740
Male	47.8	46.5	1.3	0.201
Eligible for free/reduced-price lunch	83.5	84.0	−0.6	0.537
English language learner	6.9	7.0	−0.1	0.874
Special education	5.4	5.8	−0.4	0.476
Overage for eighth grade	18.4	19.3	−0.9	0.281
Eighth-grade reading proficiency				
Did not meet standards (level 1)	7.4	6.7	0.7	0.183
Partially met standards (level 2)	62.4	63.5	−1.1	0.283
Fully met standards (level 3)	29.3	29.1	0.2	0.842
Met standards with distinction (level 4)	1.0	0.8	0.2	0.396
Eighth-grade math proficiency				
Did not meet standards (level 1)	20.2	19.9	0.3	0.763
Partially met standards (level 2)	47.6	46.8	0.8	0.444
Fully met standards (level 3)	30.5	31.6	−1.1	0.239
Met standards with distinction (level 4)	1.7	1.7	0.0	0.916
Missing eighth-grade pretests				
Math proficiency	0.8	1.2	−0.4	0.097
Reading proficiency	3.3	3.3	0.0	0.998
Sample size (total = 14,969)	6,230	8,739		
Number of lotteries (total = 199)				
Number of SSCs (total= 84)				

Sources: Findings are based on data from the NYCDOE HSAPS, NYCDOE state test data for eighth graders in 2004–05 (cohort 1), 2005–06 (cohort 2), and 2006–07 (cohort 3) plus NYCDOE enrollment files for the 2005–06 to 2010–11 school years.

Notes: A two-tailed *t*-test was applied to the estimated difference. Statistical significance levels are indicated as ** = 1 percent; * = 5 percent.

A chi-square test was used to assess the statistical significance of the overall difference between lottery winners and control group members reflected by the full set of baseline characteristics in the table. The resulting chi-square value is not statistically significant (*P*-value = 0.374).

information for the study's final lottery sample. Treatment- and control-group comparisons in these tables were constructed in a way that accounts for the fact that randomization of sample members was blocked by lottery. To do so, the treatment- and control-group difference in the mean value of each student baseline characteristic (X) was estimated as the value of β_X from the following regression:

$$X_i = \sum_{j=1}^{J} \alpha_j \cdot I_{ji} + \beta_X \cdot T_i + v_i \tag{B.1}$$

where

X_i = the value of baseline characteristic X for sample member i,
I_{ji} = 1 if sample member i was in the analysis for SSC lottery j and 0 otherwise,
T_i = 1 if sample member i won his SSC analysis lottery and 0 otherwise,

v_i = a random error that is independently and identically distributed across sample members within lotteries.

The first column in Tables 1 and B.1 reports sample mean values of each baseline characteristic for lottery winners. The third column reports estimates of the treatment- and control-group difference in means for each baseline characteristic ($\hat{\beta}_X$). The fourth column reports the level of statistical significance (P-value) for each estimated difference of means. The second column reports the inferred mean of each baseline characteristic for control group members. This was obtained by subtracting the estimated treatment- and control-group difference of means from the estimated mean for treatment-group members. As can be seen from the table, there is no discernible baseline difference between treatment- and control-group members in our final lottery sample.

APPENDIX C

Estimating Mean Outcomes for Target SSC Enrollees and Their Control Group Counterparts

Each table that reports estimates of the effect of enrolling in an SSC on a student outcome presents these findings in the context of corresponding mean outcomes for target SSC enrollees and their control group counterparts. Because these two types of students cannot be identified individually, their mean outcomes cannot be observed directly. Instead they must be inferred. This is made possible by the statistical properties of randomization plus two plausible assumptions. Figure C.1 presents the conceptual model that underlies this inferential process. The model portrays two main subgroups of SSC lottery winners: those who enroll in an SSC (a large majority) and those who do not enroll in an SSC (a small minority who do not

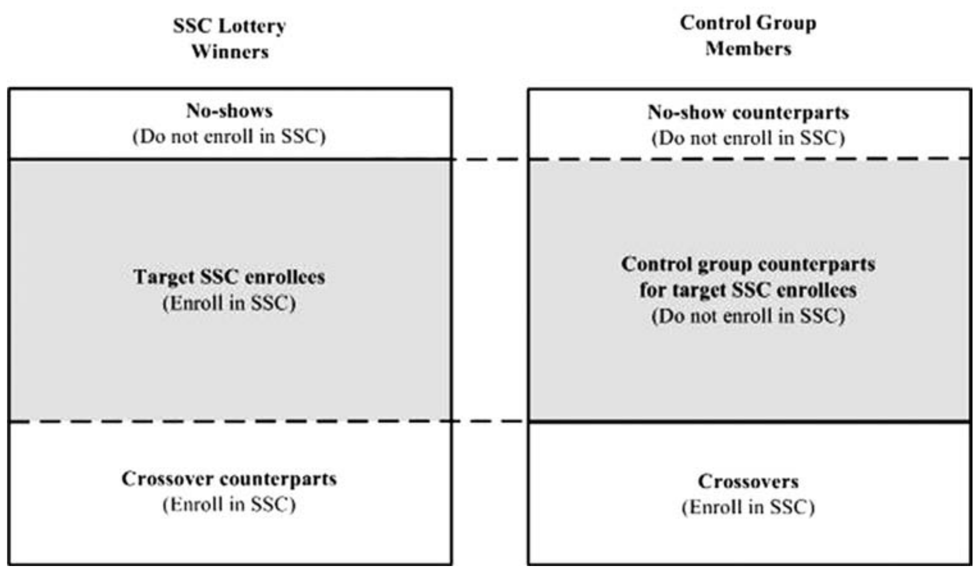


Figure C.1. Model of SSC Enrollment Among Lottery Winners and Control Group Members.

enroll and thus become no-shows).[39] The model also portrays two main subgroups of SSC control group members: those who do not enroll in an SSC (a large majority) and those who do enroll in an SSC (a small minority who become crossovers).[40]

Assume (as seems plausible) that students are more likely to enroll in an SSC if they win an SSC lottery than if they do not win. Given this assumption, which is often referred to as *monotonicity* or the absence of *defiers* (Angrist, Imbens, & Rubin, 1996), there are two more subgroups in the model: *no-show counterparts* among control group members (control group members who do not enroll in an SSC and would not have done so if they had won their SSC lottery) and crossover counterparts among SSC lottery winners (lottery winners who enroll in an SSC and would also have done so if they had been randomized to the control group).[41] Lottery winners who are neither no-shows nor crossover counterparts are target SSC lottery enrollees. This is the subgroup of students for whom SSC effects were estimated and it has a counterpart subgroup among control group members. Because these students complied with their first SSC lottery, they would be referred to as compliers in the literature.

Randomization ensures that in expectation (1) the proportion of SSC lottery winners who are no-shows (P_{NS}) equals the proportion of control group members who are no-show counterparts and (2) the proportion of control group members who are crossovers (P_{CO}) equals the proportion of SSC lottery winners who are crossover counterparts. Hence, the proportion of target SSC enrollees among SSC lottery winners equals the proportion of control group members who are their counterparts.

Now assume that winning an SSC lottery per se has no appreciable direct effect on future student academic performance and that only by causing students to enroll in an SSC can winning a lottery affect these outcomes (which often is referred to as an *exclusion* restriction, Angrist, Imbens, & Rubin, 1996). This assumption is highly plausible in the present situation because students do not even know they are in SSC lotteries. Consequently, randomization ensures that in expectation, mean outcomes for crossovers in the control group ($\bar{Y}_{CO}$) equal those for crossover counterparts among SSC lottery winners.

Now, note that the mean value of an outcome for all SSC lottery winners who enroll in an SSC ($\overline{Y}_{LWE}$) is a weighted average of mean outcomes for target SSC enrollees ($\overline{Y}_{tarE}$) and crossover counterparts with weights equal to the relative size of each group. Then, note that the observed mean outcome for crossovers ($\hat{Y}_C$) is an unbiased estimate of the mean outcome for crossover counterparts. Together, these facts imply that

$$\overline{Y}_{LWE} = \left[\frac{1 - P_{NS} - P_{CO}}{1 - P_{NS}}\right] \cdot \overline{Y}_{tarE} + \left[\frac{P_{CO}}{1 - P_{NS}}\right] \cdot \overline{Y}_{CO}. \quad (C.1)$$

Solving equation C.1 for $\bar{Y}_{tarE}$ yields

$$\overline{Y}_{tarE} = \left[\frac{1 - P_{NS}}{1 - P_{NS} - P_{CO}}\right] \cdot \overline{Y}_{LWE} - \left[\frac{P_{CO}}{1 - P_{NS} - P_{CO}}\right] \cdot \overline{Y}_{CO}. \quad (C.2)$$

[39] See Bloom (1984) for a discussion of no-shows.
[40] See Bloom et al. (1997) for a discussion of crossovers.
[41] Angrist, Imbens, and Rubin (1996) refer to no-shows and their control group counterparts as “never takers” and to crossovers and their treatment-group counterparts as “always takers.”

In this way, the implied value of $\overline{Y}_{tarE}$ can be inferred from observed values of P_{NS}, P_{CO}, $\overline{Y}_{LWE}$, and $\overline{Y}_{CO}$. This is how the findings in column one of Tables 2 to 7 in the paper were obtained. Findings in column two of the tables, which are estimates of the mean outcome for control group counterparts, are obtained by subtracting the estimated effect of enrolling in an SSC in column three from the estimated mean outcome for target SSC enrollees in column one.

REFERENCES

Angrist, J. D., Imbens, G. W., & Rubin, D. (1996). Identification of casual effects using instrumental variables (with discussion). Journal of the American Statistical Association, 91, 444–472.

Bloom, H. S. (1984). Accounting for no-shows in experimental evaluation designs. Evaluation Review, 8, 225–246.

Bloom, H. S., Orr, L. L., Bell, S. H., Cave, G., Doolittle, F., Lin, W., & Bos, J. M. (1997). The benefits and costs of JTPA title II-A programs: Key findings from the National Job Training Partnership Act Study. The Journal of Human Resources 32, 549–576.

APPENDIX D

Comparing Distributions of New York State Regents Scores for Sample Members Who Did and Did Not Attended SSCs

Figures D.1 to D.4 compare the distributions of New York State Regents scores in English language arts, science, global history, and American history for members of the four-year follow-up sample who enrolled in an SSC with corresponding distributions for sample members who enrolled in some other NYCDOE high school. The figures illustrate that there is no difference between SSCs and other NYCDOE high schools in their teacher influence on the percentage of students who just pass these examinations (i.e., score at or just above 65 points).

APPENDIX E

Estimates of the Average Effects of Winning an SSC Lottery on Four-Year High School Graduation and College Readiness

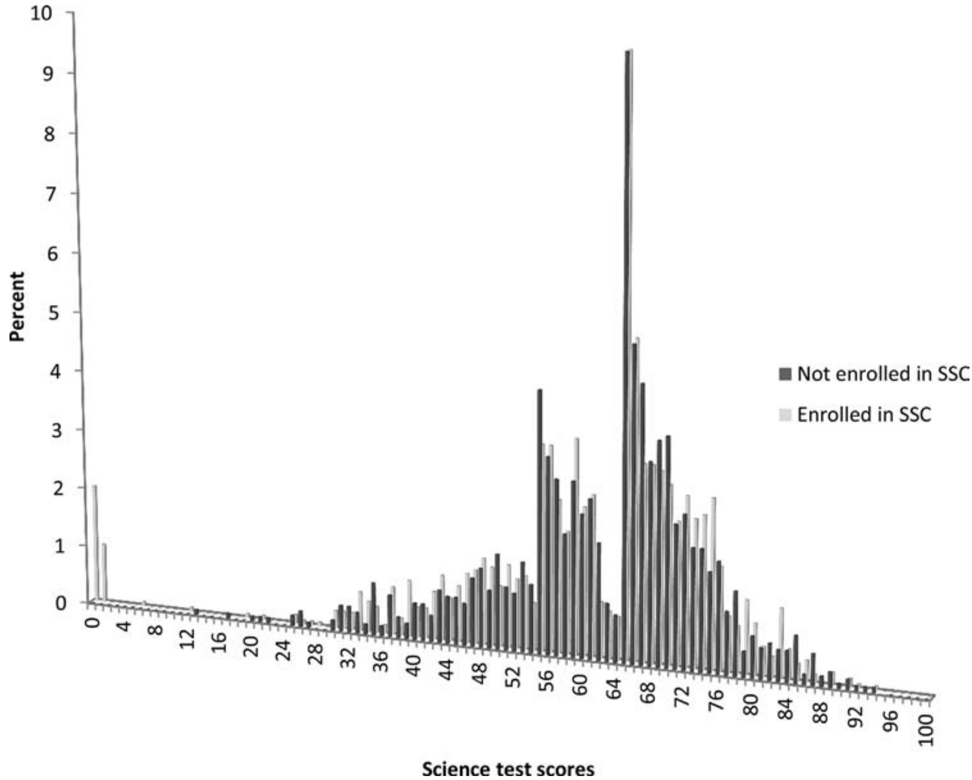
Table E.1 reports estimates of the average effect of winning a student's first SSC lottery on high school graduation and college readiness. These findings represent the average effect of intent-to-treat and are the direct experimental counterparts to the local average treatment effects reported in Table 4.

APPENDIX F

Estimating SSC Effects on High School Cost[42]

The present cost analysis uses data on annual *direct service expenditures per pupil* obtained from School-Based Expenditure Reports for the 2005–06 through 2010–11

[42] This appendix was written in collaboration with Professor Robert Bifulco from the Maxwell School of Citizenship and Public Affairs at Syracuse University. Further details and a more extensive analysis are presented in Bifulco, Bloom, and Unterman (2014).

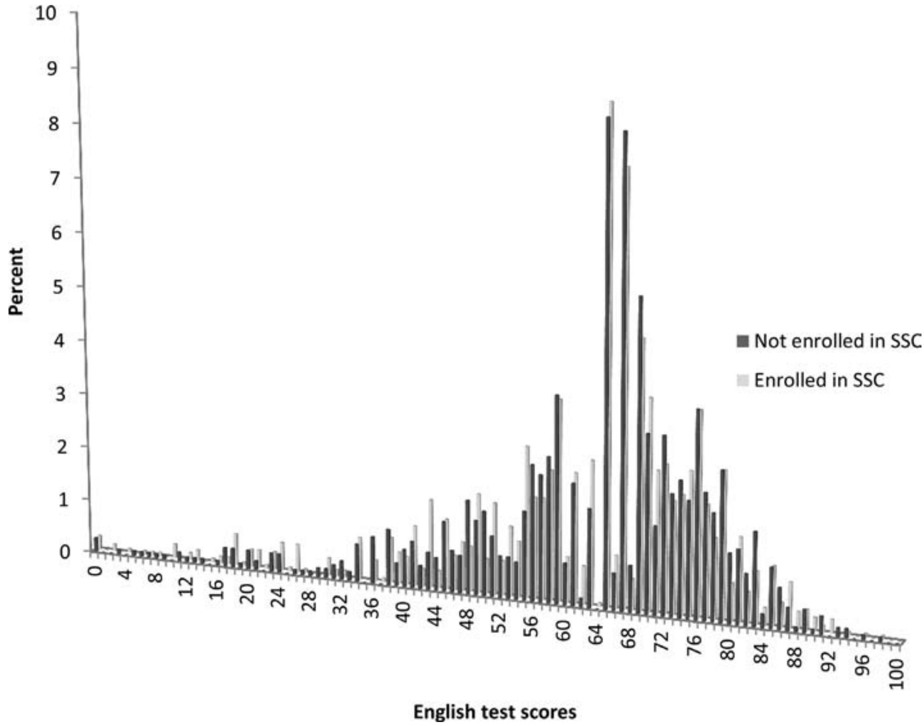


Sources: Findings are based on data from the NYCDOE HSAPS for eighth graders in 2004–05 (cohort 1), 2005–06 (cohort 2), and 2006–07 (cohort 3), plus data from NYCDOE files on student attendance, course credits, Regents examination scores, administrative transactions, and enrollment for the 2005–06 to 2010–11 school years.

Figure D.1. Distribution of Scores on the New York State English Regents Examination for SSC Enrollees and Non-SSC Enrollees in the Four-Year Follow-Up Sample.

school years prepared by the NYCDOE. Direct service expenditures include all expenditures for services provided directly to students mainly in the school building during the school day and year. These expenditures include those recorded in each school's budget plus district-level expenditures for services provided to each school or its students, including food, transportation, and building services. New York City's school-based expenditure reporting system makes extensive efforts to allocate all district-level expenditures to individual schools. Any spending other than direct-service expenditures are allocated to schools on a per pupil basis and thus do not vary across schools on a per pupil basis. Therefore, direct service expenditures per pupil is the most comprehensive measure available for comparing annual school operating costs.

The majority of expenditures reported in the NYCDOE School-Based Expenditure Reports are for salaries of teachers and other school staff members. Salaries for most school staff members are determined by collective bargaining agreements and will vary for individuals in the same position depending on their teaching experience and education level. Thus, even between schools with the same number of teachers and other staff members, salary expenditures can differ appreciably. If

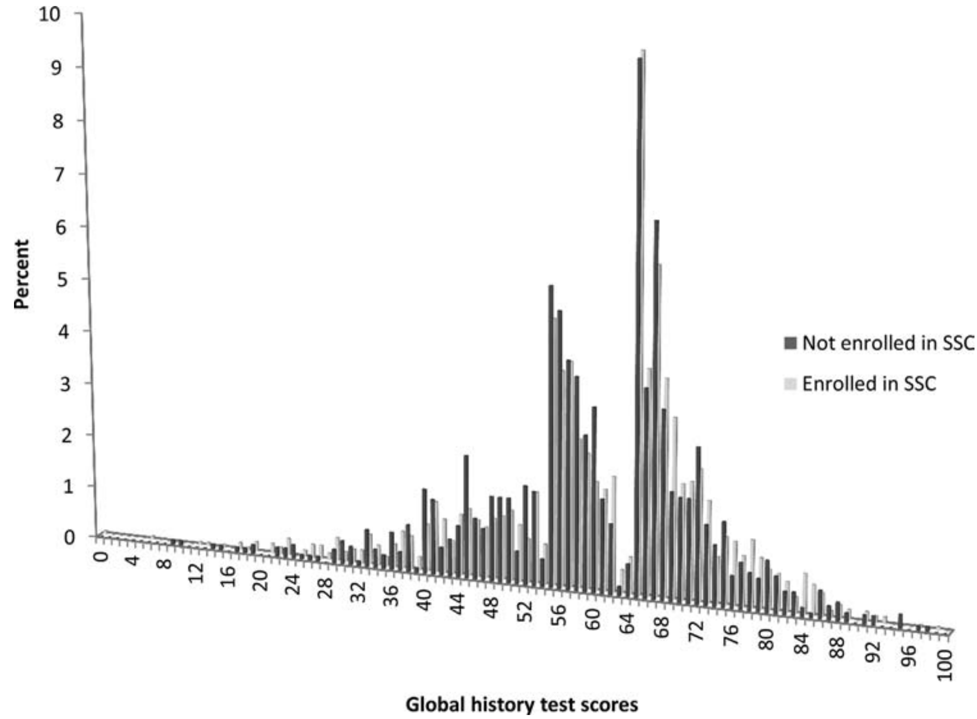


Sources: Findings are based on data from the NYCDOE HSAPS for eighth graders in 2004–05 (cohort 1), 2005–06 (cohort 2), and 2006–07 (cohort 3), plus data from NYCDOE files on student attendance, course credits, Regents examination scores, administrative transactions, and enrollment for the 2005–06 to 2010–11 school years.

Figure D.2. Distribution of Scores on the New York State Science Regents Examination for SSC Enrollees and Non-SSC Enrollees in the Four-Year Follow-Up Sample.

salaries accurately reflect the value that school personnel have in their best alternative use, then salary expenditures (plus benefits) are an appropriate measure of the cost of these personnel. However, there is good reason to question whether salaries accurately reflect the relative value of different teachers and school staff members. For instance, a large literature on teacher effectiveness suggests that experience and education levels, which account for virtually all of the variation in teacher salaries, are only weakly related to teacher quality (Clotfelter, Ladd, & Vigdor, 2006; Goldhaber, 2008; Goldhaber & Brewer, 1997; Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004). Thus, because SSCs tend to employ less experienced and less credentialed teachers who receive lower salaries than teachers in other high schools, per pupil spending in SSCs may understate their true costs relative to those of other high schools.

To remove this source of variation from our cross-school cost comparisons, we created a measure of *adjusted direct service expenditures per pupil*, which values the time of all teachers at all high schools for a given school year according to the district-wide average high school teacher salary for that year. Direct service expenditures per pupil for each school for a given school year were computed as follows.

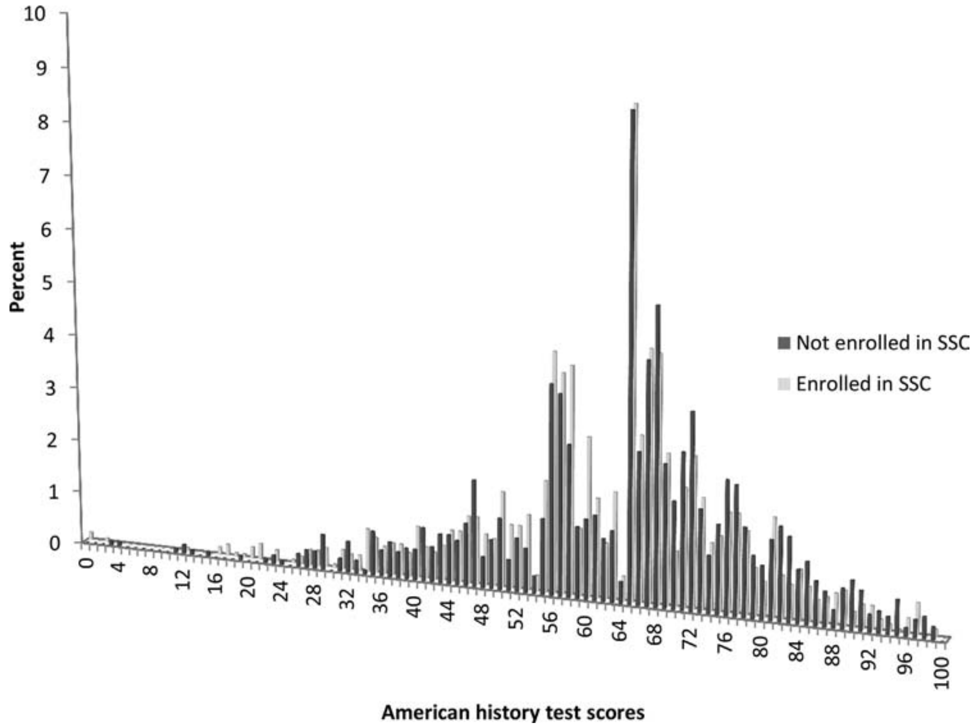


Sources: Findings are based on data from the NYCDOE HSAPS for eighth graders in 2004–05 (cohort 1), 2005–06 (cohort 2), and 2006–07 (cohort 3), plus data from NYCDOE files on student attendance, course credits, Regents examination scores, administrative transactions, and enrollment for the 2005–06 to 2010–11 school years.

Figure D.3. Distribution of Scores on the New York State Global History Regents Examination for SSC Enrollees and Non-SSC Enrollees in the Four-Year Follow-Up Sample.

1. Using data from the New York State Personnel Master File for each NYCDOE high school in our analysis,[43] we computed annual average teacher salaries for each school in our analysis separately and for all teachers from all of the schools combined. We refer to the latter as the *district-wide* annual average teacher salary.
2. We then used information from the Personnel Master File (which is part of the New York State Education Department's Basic Education Data System) to determine the total number of teachers at each high school. Using these teacher counts, we created two estimates of total teacher salaries for each school. The first measure equaled the product of the number of teachers at the school times our estimate of the district-wide mean teacher salary. The second measure equaled the product of the number of teachers at the school times our estimate of the mean teacher salary for that school.

[43] This set of high schools includes all NYCDOE schools serving grades 9 to 12 (or some subset of grades between 9 and 12), except small schools of choice not included in the sample used to estimate SSC effects on high school graduation, schools that closed between 2005–06 and 2010–11, schools that exclusively serve special populations including residential programs and night schools, and schools that primarily serve grades other than 9 to 12.



Sources: Findings are based on data from the NYCDOE HSAPS for eighth graders in 2004–05 (cohort 1), 2005–06 (cohort 2), and 2006–07 (cohort 3), plus data from NYCDOE files on student attendance, course credits, Regents examination scores, administrative transactions, and enrollment for the 2005–06 to 2010–11 school years.

Figure D.4. Distribution of Scores on the New York State American History Regents Examination for SSC and Non-SSC Enrollees in the Four-Year Follow-Up Sample.

3. We then computed the difference between the first and second teacher salary measures for each school and added it to the school's total direct student service expenditures. For schools with teacher salaries that were below the district-wide mean (like most SSCs), this adjustment increased reported costs. For schools with teacher salaries that were above the district-wide mean (like many schools attended by control group members), this adjustment reduced reported costs.
4. Lastly, we divided the adjusted total direct student service expenditures for each school by the number of students enrolled in it. This created our adjusted measure of direct service expenditures per pupil for each school.

The preceding steps produced an unadjusted and an adjusted measure of direct service expenditures for each NYCDOE high school in our analysis for each school year from 2005–06 through 2010–11. This information was then linked to students in our study sample based on the NYCDOE high school they were enrolled in during October of each of their five follow-up years (October is the official New York state date for determining student enrollment). For sample members who were identified as a school dropout based on NYCDOE discharge codes, a direct service cost of zero was entered for each year following the last year in which they were officially enrolled in high school. For students who were known to have graduated from high

Table E.1. Estimated effects of winning a student's first SSC lottery on four-year high school graduation and college readiness: cohorts 1 and 3.

Outcome (%)	Lottery winners	Control group counter-parts	Estimated effect	*P*-value for estimated effect
Graduation				
Graduated from high school	69.6	63.6	6.0**	0.000
Local diploma granted	17.1	15.5	1.6	0.061
Regents diploma granted	44.9	41.3	3.5**	0.002
Advanced Regents diploma granted	7.5	6.8	0.8	0.172
College readiness				
English Regents examination score of 75 or above	39.5	35.4	4.1**	0.000
Math A Regents examination score of 75 or above	24.3	24.6	−0.3	0.790
Sample size	5,020	7,110		

Sources: Findings are based on data from the NYCDOE HSAPS for eighth graders in 2004–05 (cohort 1), 2005–06 (cohort 2), and 2006–07 (cohort 3), plus data from NYCDOE files on student attendance, course credits, Regents examination scores, administrative transactions, and enrollment for the 2005–06 to 2010–11 school years.

Note: A two-tailed *t*-test was applied to the estimated effect. Statistical significance levels are indicated as ** = 1 percent; * = 5 percent.

school during their fourth follow-up year, a value of zero was entered for their direct service cost in their fifth year.

This left 4.8, 4.8, 4.5, and 6.3 percent of our sample members who were missing cost data for their first, second, third, or fourth follow-up years, respectively, plus 15.7 percent who were missing cost data for their fifth follow-up year. Values for these missing data were imputed using a single replication of the multiple-imputation model that was used for all other missing data. Thus, findings in Table 9 are for all members of the four-year follow-up sample. Corresponding findings (which are not reported) that omit members of the four-year follow-up sample with missing cost data were very similar.

REFERENCES

Bifulco, R., Bloom, H. S., & Unterman, R. (2014). Estimating the effects of New York City's small high schools of choice on students' high-school costs. Working paper. New York, NY: MDRC.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. NBER Working Papers 11936. Cambridge, MA: The National Bureau of Economic Research.

Goldhaber, D. (2008). Teachers matter, but effective teacher quality policies are elusive. In H. F. Ladd & E. B. Fiske (Eds.), Handbook of research in education finance and policy (pp. 146–165). New York, NY: Routledge.

Goldhaber, D., & Brewer, D. (1997). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. Journal of Human Resources, 32, 505–523.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. Econometrica 73, 417–458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. American Economic Review, 94, 247–252.