Economics of Education Review 28 (2009) 585–599

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Mission matters: The cost of small high schools revisited

Leanna Stiefel[a,*], Amy Ellen Schwartz[b], Patrice Iatarola[c], Colin C. Chellman[d]

[a] Wagner Graduate School, New York University, 298 Lafayette Street, 3098, New York, NY 10012, United States
[b] Wagner and Steinhardt Schools, Institute for Education and Social Policy, New York University, 298 Lafayette Street, New York, NY 10012, United States
[c] Florida State University, 1209G Stone Building, Tallahassee, FL 32306, United States
[d] The City University of New York, 555 West 57th Street, Room 1150, New York, NY 10019, United States

ARTICLE INFO

Article history:
Received 5 March 2008
Accepted 21 January 2009

JEL classification:
H40
I20
I22

Keywords:
Costs
Efficiency
Educational economics
Resource allocation

ABSTRACT

With the financial support of several large foundations and the federal government, creating small schools has become a prominent high school reform strategy in many large American cities. While some research supports this strategy, little research assesses the relative costs of these smaller schools. We use data on over 200 New York City high schools, from 1996 through 2003, to estimate school cost functions relating per pupil expenditures to school size, controlling for school output and quality, student characteristics, and school organization.

We find that the structure of costs differs across schools depending upon mission—comprehensive or themed. At their current levels of outputs, themed schools minimize per pupil costs at smaller enrollments than comprehensive schools, but these optimally sized themed schools also cost more per pupil than optimally sized comprehensive schools. We also find that both themed and comprehensive high schools at actual sizes are smaller than their optimal sizes.

1. Introduction

Reducing the size of high schools is a strategy popular among policymakers and philanthropists hoping to improve outcomes for students in urban school districts. In 2000, for example, the Oakland Unified School District began to develop 10 new small high schools and Chicago opened 23 small high schools over the past few years (Gewertz, 2006; Oakland Unified School District, 2000). Additionally, in 2001, seven more districts across the country committed to creating small high schools and restructuring larger ones into multiple small schools with support from the Carnegie Corporation and the Gates Foundation.[1] The small school reform model is clearly a popular solution to poor high school performance in large urban districts (Iatarola, Schwartz, Stiefel, & Chellman, 2008).

At the same time, the substantial enthusiasm for small schools is not matched by a similarly substantial body of research demonstrating the effectiveness of reducing school size.[2] While some research points to higher outcomes, it is not clear whether this is due to size or greater resources. Such research becomes increasingly important as the number of small high schools grows and as ques-

* Corresponding author. Tel.: +1 212 998 7437; fax: +1 212 995 4162.
E-mail address: leanna.stiefel@nyu.edu (L. Stiefel).

[1] The initiative was named "New Schools for a New Society." Six districts (Boston, Chattanooga, Providence, Sacramento, San Diego and Worcester) each received $8 million and a seventh (Houston) received $12 million. The Carnegie Corporation and the Gates Foundation were joined by the Open Society Institute in spearheading a similar initiative in New York City – *New Century High Schools* – that was funded in 2001 at $30 million over five years (Carnegie Corporation of New York, 2006).

[2] The definition of "small" varies significantly across the policy, advocacy/foundation, and empirical research literature. See Appendix Table A.1 for the definitions of small, medium and large enrollment sizes referenced in prominent examples from the literature.

0272-7757/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.econedurev.2009.01.005

tions emerge about whether small schools can deliver improved student test scores better than larger high schools or, more fundamentally, whether and how size matters at all (Bloomfield, 2006; Ravitch, 2006; Schneider, Wyse, & Keesler, 2007; Viadero, 2006). More importantly, as philanthropic resources decline, when policymakers allocate scarce public resources to diverse and numerous needs, we need to understand whether increasing the number of small schools is an effective use of those resources (Arenson, 2002). Do smaller schools cost more or less? Is decreasing school size enough, or are other changes needed to transform failing comprehensive high schools into effective small schools? This paper continues a long tradition in economics of estimating school cost functions (Cohn, 1968; Riew, 1966) and contributes to current research on the relationship between high school costs and size by examining the cost of schools using data on public high schools in New York City (NYC).

NYC is an especially good place to study the cost-effectiveness of small high schools due, in large part, to its long-lived and vibrant small schools movement. Dating back to the 1960s, the first wave of small school reform focused on providing alternatives for students who were not succeeding in large and traditionally oriented schools. The second wave of reform emerged in the mid-1990s, expanding the purview of small schools to a more academically oriented focus (Stiefel, Berne, Iatarola, & Fruchter, 2000). The third and current wave of reform, by far the largest, promises to transform the landscape of secondary education in NYC, nearly doubling the number of high schools through the creation of small schools (Iatarola et al., 2008). Equally important, NYC has unusually rich data on its high schools including information on school expenditures, outputs, size, and characteristics of students, allowing us to construct a panel data set on over 200 high schools from 1995–1996 to 2002–2003 (hereafter 1996–2003).

This study improves on previous research examining the relationship between school costs and school size in six key ways. First, we use a larger sample of high schools than has previously been used, providing more statistical power. Second, the range in school size is considerably broader than in previous studies, ranging from fewer than 300 students to over 2000 students. Third, while previous research used only a single output measure – the graduation rate, achievement test scores, or pass rates – we include multiple measures of performance. Fourth, while previous studies typically controlled for the "level" of outputs, we also include a measure of student performance prior to high school to estimate a "value-added" cost model[3]; and fifth, we employ a longitudinal rather than cross-sectional design, which allows us to use panel data methods to control for unobserved school variables. Finally, our analyses distinguish two types of high schools: *themed* schools, which range from small to medium size and provide narrowly focused curricula and course offerings, and *comprehensive* schools, which span a wide range of sizes and provide "...a variety of programs, support services and extra- and co-curricular activities [with a...] full range of required and elective courses".[4,5] While comprehensive schools are typically neighborhood-based, themed schools typically address specific academic interests of students (such as environmental studies or art) and are more loosely tied to their neighborhood. In addition, themed schools are never large. Thus, we begin to disentangle the impacts of size and mission, often conflated in previous research.

To preview our results, we find that the mission of a high school is important in determining its cost structure—estimated cost parameters capturing the elasticity of costs with respect to size differ between these types of schools. The results suggest that for themed schools, costs per pupil decline with enrollment and are at their minimum at roughly 700 students. In contrast, for comprehensive schools, costs per pupil are still dropping at 700 and continue to drop through the largest observed size (around 4000), although the marginal reductions are not substantively significant at the observed maximum. In addition, we find evidence that at current graduation rates and SAT scores, both types of high schools are, on average, too small.

Importantly, our results suggest that decreasing school size reduces costs only if small size is joined with themed mission. For example, our results indicate that transforming a 1500 student comprehensive school into two 750 student comprehensive schools will likely increase costs without improving outcomes. On the other hand, if those two small schools become themed schools, it is possible to decrease costs per pupil.

The rest of this paper is organized as follows. In Section 2, we review the literature on the costs of schools of different sizes; in Section 3, we develop our conceptual framework and in Section 4, we describe our sample, data, and variables; Section 5 presents the models and methods; in Section 6 we describe results; and in Section 7 we discuss selection of students and teachers and Section 8 concludes.

2. Previous research

The debate about the relative merits of large and small high schools dates back at least a half a century to James Conant's (1959) support for the large comprehensive high schools developed at the turn of the 20th century. Since then, research has examined a wide range of issues related to school size, such as effects on outputs and outcomes, breadth and depth of curriculum, and social aspects of schools.[6] Relatively little attention has been paid, however,

[3] To be specific, we include lagged performance-that is, performance of the students prior to entering high school-among the independent variables.

[4] NYC Department of Education High School Directory 2007–2008, Section 6 "Selecting a High School That's Right For You." <schools.nyc.gov/Offices/StudentEnroll/HSAdmissions/HSDirectory/News/SelectingAHS.htm> Retrieved November 9, 2007.

[5] Eighteen (out of 200) city high schools are vocational. Their mission and financing differ from comprehensive and themed and for that reason they are excluded from our analysis.

[6] Early empirical studies reveal a negative correlation between school size and academic outcomes, although costs are rarely included in these studies. For example, Fowler and Walberg (1991) and Fowler (1992) review literature that finds that students in small schools do better than those in large schools as measured by test scores, attendance rates, graduation or dropout rates, and participation in extracurricular activities. A more

to the relationship between school size and costs *per se* and the literature, which we review here, is fairly thin.

In an early paper, Riew (1966), using operating expenditure data on districts with a single high school, concluded that beyond an enrollment of 900, the existence of economies of scale is unclear. Cohn (1968) used district-level data on both short-run operating costs and long-run capital costs and found that the cost curve for high schools is U shaped, with the minimum cost at about 1500 students suggesting rising per pupil costs beyond 1500. Using only operating expenditures at the middle and elementary school level, Riew (1986) found declining costs in middle schools with enrollments as large as 1024, but at the elementary school level the lowest costs he found were in schools with 200–400 enrolled students. In a more recent study of Welsh students aged 11–18, Foreman-Peck and Foreman-Peck (2006) find an optimum school size of 540 pupils, with schools over 600 pupils having lower exam scores and attendance rates, controlling for prior exam scores. Colgrave and Giles' (2005) recent meta-analysis of 10 cost studies (some reviewed here) with 22 separate estimates concludes that the cost-minimizing high school size is 1540 students. All of the studies of high schools find a negative or flat relationship between size and average costs for high schools with enrollments up to around 1000 students (see Chabotar, 1989; Kumar, 1983; Watt, 1980; and Bee & Dolton, 1985 for more results and Andrews, Duncombe, & Yinger, 2002, for a review of literature on the costs of different size districts and schools).

There are several common deficiencies in the existing studies of the effect of size on school costs. First, while understanding costs requires a simultaneous consideration of the quality and quantity of output, few existing studies do so. That is, costs are likely determined by the level (or gain) in student performance produced, perhaps in multiple dimensions, as well as enrollment; costs are likely to be higher, *ceteris paribus*, when outputs are higher. The empirical literature on costs often includes only one output measure and sometimes none at all. Studies specifically relating costs to outputs suggest that small schools can be cost-effective. Stiefel et al. (2000) found costs of small high schools (less than 600 students) were about the same as large high schools when considered on a per graduate basis. Alternatively, Kuziemko (2006) examined shocks in enrollment in Indiana elementary schools (i.e., exogenous variation in school size) and found positive effects on math scores and attendance rates of reducing school size. Kuziemko estimated that roughly a 50% decrease in elementary school size led to a 20% increase in costs, but that the return on the additional costs yielded a net benefit of $3298.

Second, although the school is the appropriate unit of analysis for investigating school costs, district-level data are often used, largely because school-level data are unavailable (for example, Callan & Santerre, 1990; Cohn, 1968). Third, cost functions rarely incorporate a value-added (or lagged performance) measure of output (Foreman-Peck & Foreman-Peck, 2006, are a recent exception), although value-added measures are preferable from a theoretical standpoint and are, indeed, now commonplace in the production function literature. Finally, we should note that the available data measure public expenditures rather than costs. How – and to what extent – expenditures differ from costs (which, in principle, capture the minimum resources needed to produce a given output) is complicated by the presence of inefficiencies in production, among other concerns. We return to this below.

Our study improves on previous work by including several measures of outputs, including lagged output to create a value added specification, and utilizing school-level data.[7]

3. Conceptual framework

To fix ideas, a school-level education cost function captures the minimum resources needed to produce a set of constant quality outputs, given input prices and environmental and organizational features. In its general form, it can be represented as

$$c_{it} = g(q_{it}, p_{it}, n_{it}, e_{it}, o_{it}) \tag{1}$$

where c represents costs in school i at time t, q represents outputs, p represents prices of inputs, n represents enrollment, e and o are environmental and organizational features, and $g(.)$ is the functional form relating them. Note that the economic concept of cost in Eq. (1) is the minimum amount of resources (in dollars) used by each school to provide its outputs.

We measure c using expenditure data, which is a measure of resources used and note that expenditures will differ from costs if resources are not allocated in a cost-minimizing fashion. It seems unlikely that schools intend to minimize costs since they typically have neither incentive nor capacity to do so. Instead, the notion that schools act to maximize output given a budget constraint is more plausible.[8] Notice then, that if the school district acts to minimize the total cost of all schools subject to a district output constraint (for example, due to state standards), allocates resources accordingly, and schools act to maximize their own output subject to the resources given, then school resources may, in fact, reflect the minimum

recent study that uses a NCES national database concludes that when output is measured as achievement in reading and math, the relationship between high school size and performance is U shaped with the optimal size between 600 and 900 or 1200 students, depending on the subject and characteristics of the students (Lee and Smith, 1997). See Raywid (1996, 1999), Cotton (1996, 2001), Page, Layzer, Schimmenti, Bernstein, and Horst (2002) and Ready, Lee, and Welner (2004) for a broader perspective on the research on school size.

7 Although production function studies often use changes in test score based measures (gains) to capture value added, that formulation is less appealing here, where we have multiple outcomes and the "post" measures include non-test score measures such as the graduation rate while the "pre" measures are the more traditional test scores.

8 Note that the provisions of NCLB complicate the objective function for elementary and middle schools (grades 3–8), which also must meet output standards for five subgroups of students. During the time period of our data, however, the impact on high schools is minimal. In New York State, high schools are required to meet state standards that do not include subgroup performance metrics.

cost of producing the observed output.[9] Unfortunately, it is unlikely that these conditions are a fully accurate description of the objectives and actions of the various administrators, superintendents and principals in a large district like NYC. Instead, observed costs also reflect differences in efficiency, or motivation, which we aim to control for using school fixed effects, as discussed in more detail below.

Because cost functions specify output as a regressor, rather than as a dependent variable as in a production function, multiple output measures can be included, which is especially important because schools produce a variety of outputs. Environmental factors capture differences in the resources needed to educate students with different needs. For example, poor students may have less well-educated parents who are unable to help them with homework or fewer educational resources available at home such as books or computers, making the job of educating poor students more resource-intensive and hence more costly than for comparable non-poor students. In addition to student characteristics, the cost function includes features of the school organization, such as school type and, of course, size, which is our primary interest.[10]

In this formulation, input prices and output quantities are exogenous. However, the prices of inputs (for example teacher salaries) are determined by *district* administrators rather than *school* administrators and do not vary across schools within a district. That is, the schedule of salaries is the same for all schools and, although there may be some differences in the prices of other inputs, within-district differences in prices are likely to be small and relatively unimportant in the school budget. Salaries can, and do, change over time as new contract terms apply and we capture these effects using year dummies.

Economic theory suggests several possible reasons that costs (measured per pupil) might vary with school size or, put differently, that there may be economies or diseconomies of scale in schooling. To begin, costs might decline with size as the fixed costs of schools are spread over a larger student body. Fixed costs might include some aspects of the physical plant such as the gymnasium or cafeterias as well as personnel, such as principals, or other inputs, such as library books. The key is that these inputs are, in some measure, indivisible. Alternatively, economies of scale may reflect the gains from specialization and division of labor, for example, if teachers are able to specialize in offering courses in their areas of expertise. On the other hand, diseconomies of scale may set in at some point, due to limits to the ability of principals (or districts) to manage a large school or to congestion in the utilization of fixed resources. Which of these effects dominates is, essentially, an empirical matter.

Estimating this cost function requires data on school spending, student outcomes, and student characteristics, as well as information on environmental/neighborhood characteristics and organizational features of each school. In the absence of these data, school and borough fixed effects can be used to capture unobserved time-invariant characteristics of schools, their neighborhoods, organizational structure, and, potentially unobserved efficiency.

4. Data, sample, and measures

NYC provides a unique opportunity for studying small schools through the richness of data available as well as the scope of the small schools movement in the city. As the largest school system in the country, and as part of the early 1990s small schools movement, NYC high schools offer a wider range of school sizes and student types than available elsewhere and over a longer period than used in other small schools cost studies. Data on school spending, student outcomes, and student and teacher characteristics are available for almost a decade, providing more statistical power for our analysis as well as more appropriate measures for a cost analysis.

4.1. Data

We obtained data from the NYC Department of Education (NYCDOE) based on school-based expenditure reports, cohort graduation reports, and high school report cards. *School-Based Expenditure Reports* (SBER) provide per pupil expenditure information (for general education students), disaggregated into several functions and programs, for each high school for each year (see Appendix A for more detail) and, critical to this study, general education student enrollment.[11] The *cohort report* data track students for 4 years from the point they enter the 9th grade, accounting for dropouts, transfers in and out of the school, and entrants to and withdrawals from the system. *Annual School Reports* (ASRs) provide student demographic and test score data, teacher characteristics such as descriptions of teachers (e.g., percent with Masters' degrees) and results of student performance on several Regents' examinations, the New York State tests mandated for graduation. Below we provide detail on our sample of schools and important variables in our models.

4.2. Sample

This study uses a panel of data for a set of high schools serving 10th through 12th graders for the 1996 through 2003 graduating classes. In 2003, over 220 schools served nearly 300,000 high school students in New York City. We exclude specialized (magnet) high schools, vocational, "last chance," or general equivalency diploma programs (GED).[12] These schools have test-based admissions, do not

[9] Put differently, if the district acts to maximize output, subject to its budget constraint, allocates resources accordingly, and schools also act to maximize their output given their resources, the observed costs should be the minimum cost of producing the observed output.

[10] Despite the appeal of adjusting for differences in efficiency, no appropriate measures are available at the school level. To some extent, differences in efficiency will be captured by school fixed effects, discussed below.

[11] The expenditure data were published for the first time in a comprehensive manner in 1996. Such detailed, well maintained and continuously available over an eight-year period are unique to large city school districts.

[12] "Last chance"/Transfer Alternative schools also address the needs of special populations, but students transfer into these schools in later

assess progress in a standardized manner, or use alternative assessments of student progress (different graduation diplomas and/or high school tests) and thus they mostly have considerably less data available than the other types of high schools. We excluded all citywide special education schools, which are solely attended by full-time special education students.[13] Taken together, all the excluded high schools represent approximately 24.4% of New York City high schools in 2003 but just 15.9% of high school students in that year. Put another way, in this study, we capture the majority of high school students and high schools that primarily educate general education students who are assessed in a standardized manner. Over the years 1996 through 2003, we have data on 201 individual schools and 1200 school-year observations.

4.3. Measures

We measure costs using data on *modified direct expenditures*, which represent resources used at the school level and include the inputs devoted to classroom instruction, instructional support services, leadership support, and ancillary support services (e.g., food service and school safety) but exclude building expenditures (e.g., energy and leasing costs), transportation and any of the district-level costs, such as the superintendent's office or district debt service. Thus, this measure captures resources used by the school itself to provide educational services and contains few allocated expenditures. Most importantly, building expenditures are excluded because they vary with the particular financial arrangement for the school building rather than resources devoted to learning. For example, many newer smaller schools are leased rather than owned outright; many older schools are fully depreciated and require more in energy and maintenance than newer buildings, but these added expenditures do not contribute to the educational function. Transportation expenditures are excluded as well because they capture neighborhood and geographic characteristics that are largely a function of whether a school is walking distance, accessible by subway or serves full-time special education students.[14] Although the expenditure numbers cannot be regarded as perfectly capturing school costs of production, they are among the best school-level resource figures in the nation and offer the opportunity to shed light on the relationship between costs and school size. That said, while we use the terms costs and expenditures interchangeably in what follows, the underlying measures are based upon reported expenditures.[15]

We measure school size as the enrollment of general education students, including those receiving part-time special education and other services. Following both the research literature and recommendations of school policymakers and advocates, we also distinguish five size categories, specifically very small (30–300 students), small (301–500), medium (501–1200), large (1201–2000) and very large (2001 and larger). Appendix Table A.1 more fully describes the rationale for these cut points.

Although often ignored in empirical work, high schools in NYC, as elsewhere, span a wide range of missions and goals, with corresponding differences in curricula, students and, in some measure, organizational and administrative structures within the Department of Education. We distinguish two non-overlapping categories of high schools—comprehensive and themed. *Comprehensive* high schools are traditional high schools that are typically neighborhood-based. They generally accept students based on geographic area of residence and do not require an entrance exam or audition. They usually provide several levels of difficulty for courses and multiple choices of languages and so on. Seventy-eight percent of the observations in our study belong to this category.[16] *Themed* schools address the specific academic interests of students through special academic programs or themes, such as law, business, and health care, among others. This category includes the city's official "small" schools—both, those formed over a decade ago in the first wave and new ones that are part of the second and current waves. The themed schools in total account for 22% of the observations in the sample.

We use multiple measures of high school outputs including average math and verbal scores on SAT examinations and SAT test-taking rates, in addition to the traditional 4-year cohort graduation rates.[17] Further, we include the percent of entering 9th grade students passing the state 8th grade math exam, which captures performance prior to entering high school. While including multiple outputs is realistic, the underlying production process is probably

grades, e.g., incarcerated students and pregnant teenagers. The General Education Diploma (GED) programs' goal is for its enrolled students to attain a GED in situations where an alternative to a traditional high school diploma is necessary.

[13] Citywide special education schools are those intended for student with severe disabilities. 5.7% (5.6%) percent of full-time special education students attend regular high schools over all the years in our sample (in 2003), but we do not include their expenditures and we do not include them in enrollment numbers.

[14] Alternative expenditure measures, including total and instructional, yield similar results and thus we do not show results using these other measures. The specific components of each measure are listed in Appendix A.

[15] The largest excluded categories of costs are building expenses (maintenance, energy and depreciation), due, in large part, to the unavailability of the necessary data to adequately control for differences in these costs across schools. It is, of course, possible that building cost could vary with size, but the sign and magnitude of the relationship is unclear *a priori*. Note, also, that the size of the school does not dictate the size of the school building. Large schools may make use of more than one building. Small schools may be housed together in a single large building. Similarly, transportation costs may vary with size to the extent that size is correlated with the students' commuting distances. As an example, if students in smaller schools have shorter commutes, then costs will be lower relative to larger schools. Alternatively, if smaller schools have greater differentiation in school characteristics, suggesting students travel farther to school, costs would be higher. These are important empirical questions that are outside the scope of this paper.

[16] We report school-year observations for each type here, which are numbers of schools by type by year multiplied by the number of years in which a school appears.

[17] Other output data, such as passing rates on Regents' exams (English and math) are available. We do not include these additional measures of outputs because of their high correlation with graduation rates and SAT scores and the likelihood that their inclusion would further limit the identification of independent relationships of the output variables to spending.

Table 1
Descriptive statistics for model variables, all schools in 2003 ($N = 180$).

	Obs	Mean	Std. dev.	Min	Max
Dependent variables					
Cost per pupil	180	$8205	$1763	$5744	$19,857
Enrollment[a]	180	1283	1152	92	4,204
Type (categorical)					
Comprehensive	180	0.77	0.42	0.00	1.00
Themed	180	0.23	0.42	0.00	1.00
Size (categorical)					
Very small (<300)	180	0.11	0.31	0.00	1.00
Small (301–500)	180	0.24	0.43	0.00	1.00
Medium (501–1200)	180	0.29	0.45	0.00	1.00
Large (1201–2000)	180	0.09	0.29	0.00	1.00
Very large (2001+)	180	0.28	0.45	0.00	1.00
Outputs					
% graduate, 4-year cohort	180	57.91	21.33	1.10	100.00
SAT Math and Verbal	178	836.74	97.92	749.00	1163.00
% taking SAT	154	36.61	12.39	2.10	65.20
Student characteristics					
% cohort passing Math, 8th grade	154	17.07	17.34	0.00	98.90
% Asian	158	9.78	12.71	0.00	77.40
% black	158	40.17	26.75	0.30	93.00
% Hispanic	158	37.69	23.68	4.40	99.70
% white	158	12.36	17.44	0.00	82.10
% female	158	52.23	8.31	13.70	82.40
% eligible for free lunch	159	61.39	25.12	0.00	99.40
% limited English proficient	158	13.66	17.29	0.00	93.40
% recent immigrants	158	9.95	14.33	0.00	96.00
% resource room	156	6.24	3.48	0.00	25.10
Borough (categorical)					
Manhattan	180	0.33	0.47	0.00	1.00
Bronx	180	0.16	0.37	0.00	1.00
Brooklyn	180	0.29	0.46	0.00	1.00
Queens	180	0.17	0.37	0.00	1.00
Staten Island	180	0.04	0.21	0.00	1.00

Notes: Indicators for missing independent variable data (not shown) are included in regression analyses. Other years have similar patterns of missing data.
[a] General education students includes part-time special education students.

characterized by some degree of jointness in production. Put differently, the outputs may be jointly affected by changes in inputs and, as a result, independent variation is likely to be somewhat limited. As an example, increases in the SAT scores are likely related to increases in graduation rates; disentangling the separate impacts, *ceteris paribus*, will be difficult as is disentangling the independent relationships between outputs and costs will be difficult. Including multiple measures, however, will help minimize bias on estimates of the relationship between size and costs.

We include characteristics of a school's student body that have been shown to influence the costs of education such as percent of students eligible for free lunch (referred to as poor, below), limited English proficient (referred to as LEP) and receiving "resource room" services (part-time special education services), as well as the racial composition of the student body.

While our data are relatively rich, New York City is physically and administratively large and culturally diverse, with each borough, neighborhood, and school having distinct characteristics unmeasured by the other variables. We attempt to control for these differences in a number of ways. Year fixed effects control for trends in spending and common changes in variables such as managerial efficiency system-wide or changes in standards, curricula, or budgeting.[18] We begin by including borough fixed effects then estimate models using school fixed effects to capture unobserved time-invariant characteristics of schools. In this case, the size coefficients capture the direct impact of enrollment holding these unobserved time-invariant school characteristics fixed. Of course, some or all of these unobserved factors may well be due to the size itself and we might wish to include the impact of these factors in our estimates of the impact of size on costs. To do so, we also estimate our model using school random effects rather than fixed effects. Put differently, we may well wish to include the indirect effects of size as well as the direct effects, which a random effects specification allows.[19]

[18] All expenditures are in constant 2003 dollars that have been adjusted using CPI-U, Bureau of Labor Statistics. Thus, the year fixed effects do not also absorb inflationary increases in spending.

[19] In Section 6, we test whether the random effects specification is justified.

Table 2
Crosstabulation of schools by size and type, 2003.

	Comprehensive		Themed	
	Number	Percent	Number	Percent
Very small (<300)	8	5.8	11	26.8
Small (301–500)	21	15.1	22	53.7
Medium (501–1200)	44	31.7	8	19.5
Large (1201–2000)	16	11.5	0	0.0
Very large (2000+)	50	36.0	0	0.0
Total	139	100.0	41	100.0

4.4. Descriptive statistics

Table 1 displays descriptive statistics for all 180 high schools in our sample for 2003.[20] To begin, costs per pupil averaged $8205. The average high school in our sample enrolled 1283 students, and the range is large, varying from 92 to 4204 students. Twenty-eight percent of all schools are very large, 9% are large, 29% are medium sized, 24% are small and 11% are very small. In the average high school, 58% of students graduate in 4 years and, of the roughly 37% of students who take the SAT, the average math and verbal score is 837. In the average high school, just over 17% of entering 9th graders passed standardized math exams in 8th grade. The student body is 10% Asian, 40% black, 38% Hispanic, and 12% white, with slightly more female than male students (52–48%, respectively), 61% poor, 14% LEP, 10% recent immigrant, and 6% in need of resource room services (part-time special education).

As shown in Table 2, the distribution of schools across the sizes and types is quite uneven. While comprehensive high schools come in all sizes, ranging from the very small to the very large, themed high schools span a narrower range from very small to medium. Viewed from the other perspective, small and very small schools are relatively evenly split between comprehensive and themed schools; a large majority of medium schools are comprehensive and importantly, all large and very large schools are comprehensive schools. Thus, the observed variation in sizes across types is sufficient to allow us to gain some insight into the relationship between mission and size, but this variation is entirely within schools of medium size and smaller.[21]

5. Models and methods

Estimating a cost function as in (1) proceeds, then, by specifying a functional form for the cost function. Unfortunately, theory and literature provide little guidance about the functional form of a school cost function and, thus, we explore two specifications, one with dichotomous and the other with continuous size variables. In addition, we also explore the differences in the cost functions between comprehensive and themed schools using both pooled regressions, which include high schools of both types in a single regression, allowing for key coefficients to differ by type, and separate regressions, which allow for the full set of coefficients to vary by type.

Our first equation specifies size as a step function, with a series of dummy variables for size categories. This equation also includes indicators for the different high school mission types. This equation is as follows:

$$\begin{aligned}\text{lncost}_{it} = {} & \beta_0 + \beta_1\,\text{Size}_{it} + \beta_2\,\text{Themed}_{it} \\ & + \beta_3\,\text{Output}_{it} + \beta_4\text{LagOutput}_{it} \\ & + \beta_5\text{Student}_{it} + \beta_6\text{Boro}_i + \beta_7\,\text{Year}_t + \varepsilon_{it} \qquad (2)\end{aligned}$$

where lncost is the natural log of expenditure per general education pupil; size is a vector of five size dummy variables, described earlier; Themed is an indicator variable distinguishing themed schools from comprehensive high schools; Output is a vector of variables that includes the 4 year cohort graduation rate, SAT scores, and percent taking the SAT; LagOutput is the average 8th grade test score for the "graduating cohort" as they entered high school; Student is a vector of variables capturing the characteristics of the student body such as the percentage poor; Boro is a vector of fixed effects for each of NYC's five boroughs; year is a vector of time fixed effects; ε is an error term with the usual properties and i indexes schools and t indexes school years.

Our central interest lies in estimating the coefficients on the size variables, which we interpret as estimates of the differences in the per pupil cost of education across the different school types. Other variables serve to control for other determinants of the cost of education. Since costs are measured in natural logarithms, the coefficients will capture the percentage differences between these.

The second equation specifies size as a continuous variable and interacts size with mission type and final outputs:

$$\begin{aligned}\text{lncost}_{it} = {} & \beta_0 + \beta_1\,\text{lnEnroll}_{it} + \beta_2\,\text{lnEnroll}^2_{it} \\ & + \beta_3\,\text{Themed}_{it} + \beta_4\text{Themed}_{it}\,\text{lnEnroll}_{it} \\ & + \beta_5\,\text{Themed}_{it}\,\text{lnEnroll}^2_{it} + \beta_6\,\text{Output}_{it} \\ & + \beta_7\,\text{Output}_{it}\,\text{lnEnroll}_{it} + \beta_8\,\text{LagOutput}_{it} \\ & + \beta_9\,\text{Student}_{it} + \beta_{10}\,\text{Boro}_i + \beta_{11}\,\text{Year}_t + \varepsilon_{it} \qquad (3)\end{aligned}$$

where LnEnroll is the natural logarithm of the number of general students enrolled and all other variables are the same as in Eq. (2). In addition, we substitute school fixed effects (random effects) for borough fixed effects in some specifications. Here, our central concern is with estimating the relationship between costs and size and examining whether there are significant differences in the

[20] Indicators for missing data for independent variables are included in regression analyses and missing values for the independent variables have been given a value of "0", to maintain sample size, but the descriptive statistics are based on available data only. Note that our panel is unbalanced. Overall years there are 201 schools that are present for at least 2 years in the panel. In some years there are more (or fewer) and different schools.

[21] Note that NYC does have several large high schools with specialized curricula, focusing on math and science or performing arts, which demonstrate that large size does not guarantee a comprehensive mission. Because these schools admit students based on admissions exams, we exclude them from our analytic sample.

size relationship between themed and comprehensive high schools.[22]

Our final analysis estimates separate cost functions for themed and comprehensive high schools, allowing each to have its own coefficients and size relationship.

$$\begin{aligned}\text{lncost}_{it} = {} & \beta_0^{\text{T}} + \beta_1^{\text{T}}\,\text{lnEnroll}_{it} + \beta_2^{\text{T}}\,\text{lnEnroll}_{it}^2 + \beta_6^{\text{T}}\,\text{Output}_{it} \\ & + \beta_7^{\text{T}}\,\text{Output}_{it}\,\text{lnEnroll}_{it} + \beta_8^{\text{T}}\,\text{LagOutput}_{it} \\ & + \beta_9^{\text{T}}\,\text{Student}_{it} + \beta_{10}^{\text{T}}\,\text{Boro}_i + \beta_{11}^{\text{T}}\,\text{Year}_t + \varepsilon_{it} \qquad (4)\end{aligned}$$

where T = themed or comprehensive. Each of Eq. (4) is estimated using a subsample of the high schools—either the themed or the comprehensive high schools.

6. Results

In this section, we present the estimated cost functions, explore the implications of the estimates for optimal (cost minimizing) school size, and discuss possible issues of selection of students or teachers by type of school.

6.1. Cost function estimates

Table 3 presents the results of estimating cost functions using the dichotomous specifications of size as in Eq. (2). To begin, notice that the relationship between costs and the characteristics of the student body is consistent with prior research. Higher percentages of poor and of LEP students increase costs. *Ceteris paribus*, a 10% point increase in poor students raises per pupil costs by 0.5%; and such an increase in LEP and part-time special education students raises costs by 4.6% and 3.8%, respectively.[23] Coefficients on other student characteristics, while consistent with expectations, are not statistically significant at the 5% level.

The regressions indicate that, as anticipated, costs are monotonically decreasing in school size. Small schools cost almost 15% less than very small schools (the omitted category); medium size schools are even slightly cheaper; large size schools nearly 20% cheaper than very small schools and very large schools fully 28% cheaper than very small schools. Viewed somewhat differently, large schools are 5% cheaper than medium and small sized schools while very large schools are 15% cheaper than medium and small sized schools. Thus, these results point to significant economies of scale. The specification of those relationships are, however, fairly restrictive.

Interestingly, these results do not suggest that themed schools have significantly different costs from comprehensive schools, although, as shown in Table 2, mission and size are related-with themed schools relatively small and comprehensive schools all sizes. Interpreting the coefficients on outputs is more difficult, since the regressions include multiple output measures simultaneously. In principle, each coefficient captures the marginal cost of raising that single output, holding constant quantities of other outputs, and marginal costs should generally be positive. However, as noted previously, output measures are likely to move together, making precise and consistent estimation of coefficients difficult. Interpreted literally, the results suggest that a 10% point increase in 4-year cohort graduation rates reduces costs by 1.6%.

To explore these relationships more fully, we turn to a more flexible specification, measuring size using enrollment and its square and allowing the impact of size to vary between themed and comprehensive schools by including an interaction term.[24]

The estimates in Table 4 are consistent with the predictions of theory and prior evidence. Although not shown, coefficients on poor, LEP, and part-time special education students indicate that they are more costly to educate, while recent immigrants are cheaper.

As shown in Table 4, cost functions are U shaped with respect to size, with themed schools moving along a different and higher curve than comprehensive schools. That is,

Table 3
Cost regression results, dichotomous size and type variables, all schools, 1996–2003.

Variable	Coefficient (robust standard errors)
Constant	8.5764*** (0.0410)
Size categories	
Small (301–500)	−0.1468*** (0.0263)
Medium (501–1200)	−0.1512*** (0.0277)
Large (1201–2000)	−0.1984*** (0.0308)
Very large (2001+)	−0.2832*** (0.0308)
Themed	−0.0136 (0.0148)
% graduate, 4-year cohort	−0.0016*** (0.0004)
SAT score	−0.0000 (0.0001)
% taking SAT	0.0000 (0.0005)
% cohort passing Math, 8th grade	0.0006* (0.0003)
% Asian	−0.0005 (0.0008)
% black	0.0006* (0.0004)
% Hispanic	−0.0004 (0.0005)
% female	0.0010* (0.0006)
% eligible for free lunch	0.0005** (0.0002)
% limited English proficient	0.0046*** (0.0012)
% recent immigrants	−0.0032*** (0.0012)
% part-time special education	0.0038** (0.0019)
Observations	1200
R^2	0.75

Dependent variable measured in natural logs. Model includes year effects and borough effects. Indicators for missing independent variable data (not shown) are included in regression analyses. Sample description: specialized and vocational schools removed as explained in the text, leaving two types of schools: comprehensive and themed. Omitted groups: very small size category; comprehensive type category.

* Significant at 10%.

** Significant at 5%.

*** Significant at 1%.

[22] Notice that for comprehensive high schools the elasticity of costs with respect to size can be calculated as $\beta_1 + 2\beta_2 \text{lnEnroll}_{it} + \beta_7 \text{Output}_{it}$. Thus, we allow the elasticity of cost to vary with both size and output level. Calculating this elasticity for themed schools requires adding in $\beta_4 + 2\beta_5 \text{lnEnroll}_{it}$, allowing the cost elasticity to differ for themed and comprehensive schools, although the coefficients on other cost factors are held constant. The optimal size is found where this elasticity is zero, with second order conditions met.

[23] In contrast, a 10% point increase in recent immigrants lowers costs by 3.2%. This is consistent with the finding in Schwartz and Stiefel (2006) that immigrants outperform native-born students, *ceteris paribus*.

[24] An exploration of cubic terms indicated the quadratic is sufficient. Cubic terms were nearly uniformly insignificant.

Table 4
Cost regression results, continuous size and dichotomous type and size interactions, all schools, 1996–2003.

Variable	Coefficient (robust standard errors)
Constant	10.0910*** (0.5055)
ln(Size)	−0.4170*** (0.1459)
$[\ln(\text{Size})]^2$	0.0239** (0.0105)
Themed	8.0730*** (1.6820)
Themed × ln(Size)	−2.8609*** (0.5636)
Themed × $[\ln(\text{Size})]^2$	0.2503*** (0.0474)
% graduate, 4-year cohort	−0.0035** (0.0017)
% graduate, 4-year cohort × ln(Size)	0.0003 (0.0003)
SAT score	0.0007*** (0.0002)
SAT score × ln(Size)	−0.0001*** (0.0000)
% taking SAT	−0.0065* (0.0037)
% cohort passing Math, 8th grade	0.0004 (0.0003)
Student characteristic variables	Yes
Observations	1200
R^2	0.78

Dependent variables measured in natural logs. Model includes year effects and borough effects. Indicators for missing independent variable data (not shown) are included in regression analyses. Student characteristic variables included but not shown: percent Asian, black, Hispanic, female, poor, LEP, recent immigrant, and in part-time special education. Sample description: specialized and vocational schools removed as explained in the text, leaving two types of schools: comprehensive and themed. Comprehensive type is the omitted group.

* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

themed schools have higher fixed costs per pupil, with costs initially dropping more quickly (larger cost elasticity) than for comprehensive schools, then rising more rapidly. The cost of raising SAT scores is positive, although decreasing with school size, while increasing cohort graduation rates cost less, but that cost increases with size (albeit insignificantly). Costs per pupil decrease with higher percentages of students taking the SATs.

Motivated in part by the difference in the estimated economies of size in previous regressions for schools with different missions, we turn next to estimated cost functions that separate themed schools from comprehensive schools.[25] This also allows us to run a series of simulations that translate our results into a policy relevant understanding of the interplay among school size, costs, and mission.

Table 5 shows the results of separate equations for comprehensive (columns 1–3) and for themed schools (columns 4–6). Models in columns (1) and (4) include borough effects, columns (2) and (5) include school fixed effects, and columns (3) and (6) include random effects.[26] Separate Hausman tests on models in columns 3 and 6 reject the random effects specification for comprehensive schools but not for themed schools. We show all results but use coefficients from the school fixed effects regres-

[25] A Roy-Zellner test of poolability for themed and comprehensive schools (or for differences in coefficients) for models with borough effects, school fixed effects, or random effects rejects poolability (F=4.65, 4.25, 164.37 respectively and all type I error probabilities less than 0.001).

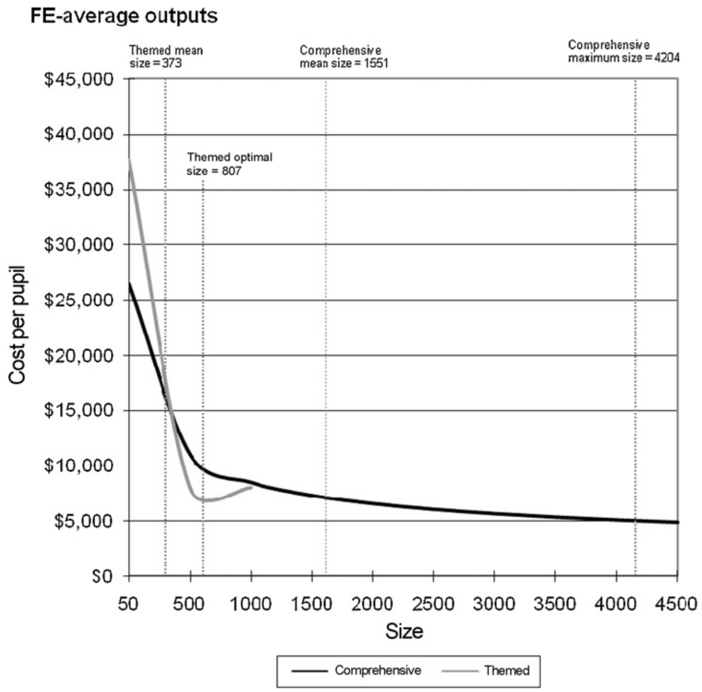
[26] Schools are entirely (and unvaryingly) nested within boroughs; using school effects precludes borough effects.

Table 5
Cost regression results for comprehensive and themed high schools, continuous size, by school type, 1996–2003.

Variable	Comprehensive			Themed		
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	9.9563*** (0.5342)	11.3642*** (0.7421)	11.2117*** (0.4231)	14.4390*** (1.8209)	17.4146*** (2.7174)	17.5197*** (1.8503)
ln(Size)	−0.3814** (0.1557)	−0.4588** (0.2186)	−0.6508*** (0.1244)	−2.1722*** (0.6238)	−2.9004*** (1.0091)	−3.1082*** (0.6380)
$\ln(\text{Size})^2$	0.0191* (0.0113)	0.0028 (0.0166)	0.0310*** (0.0093)	0.1970*** (0.0533)	0.2303** (0.0935)	0.2635*** (0.0554)
Graduation rate	−0.0046** (0.0020)	−0.0073* (0.0041)	−0.0114*** (0.0025)	0.0239*** (0.0085)	0.0145 (0.0130)	0.0187** (0.0083)
Graduation rate × ln(Size)	0.0006** (0.0003)	0.0013** (0.0006)	0.0018*** (0.0004)	−0.0043*** (0.0014)	−0.0023 (0.0022)	−0.0032** (0.0014)
SAT score	0.0002 (0.0002)	0.0003** (0.0001)	0.0004** (0.0002)	0.0004 (0.0005)	−0.0000 (0.0004)	0.0001 (0.0004)
SAT score × ln(Size)	−0.0000* (0.0000)	−0.0000** (0.0000)	−0.0001*** (0.0000)	−0.0001 (0.0001)	0.0000 (0.0001)	−0.0000 (0.0001)
% taking SAT	0.0004 (0.0006)	−0.0004 (0.0004)	0.0002 (0.0003)	0.0000 (0.0006)	0.0002 (0.0006)	0.0002 (0.0005)
% cohort passing Math, 8th grade	−0.0001 (0.0004)	−0.0006* (0.0003)	−0.0008*** (0.0003)	0.0012* (0.0007)	0.0008* (0.0005)	0.0008 (0.0006)
Student characteristic variables	Yes	Yes	Yes	Yes	Yes	Yes
Borough fixed effects	Yes	No	No	Yes	No	No
School fixed effects	No	Yes	No	No	Yes	No
Random school effects	No	No	Yes	No	No	Yes
Observations	933	933	933	267	267	267
R^2	0.81	0.94	0.76	0.74	0.89	0.65

Dependent variable measured in natural logs. Models include year effects. Robust standard errors in parentheses. Specialized and Vocational schools removed as explained in the text. Indicators for missing independent variable data (not shown) are included in regression analyses. Student characteristic variables included but not shown: percent Asian, black, Hispanic, female, poor, LEP, recent immigrant, and in part-time special education.

* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.



Note: Simulations calculated using mean student characteristics and mean outputs by school mission type – comprehensive or themed - for 2003.

Fig. 1. Simulations of high school costs for "own students" and average outputs by type, 2003.

sions in cost simulations, in the interest of clarity and ease of interpretation.

All models yield cost parameters that suggest positive fixed costs, with dropping then rising marginal costs and U shaped cost functions. Most broadly, both because of their consistency with theory and literature and their success with inclusion of school effects, we view the results in Table 5 as the best. Within each of the school types, comprehensive and themed, the specifications yield results that are broadly similar, suggesting that the bias due to unobserved differences in schools may not be problematic.[27] Coefficients on outputs, when statistically significant, demonstrate either a main effect that increases costs and a smaller, interacted effect with size that shows a decline as size increases (or the reverse). Further, when significant, coefficients on student characteristics are similar to those found in the other estimates, e.g., poor and LEP students are associated with higher costs.

6.2. *Cost simulations*[28]

The key question in this study is: do smaller schools cost less to operate or more–suggesting schools be smaller or large and, if so, how much? How does the answer depend upon the quality of the output? To gain insight into the answer, we use our estimated coefficients to simulate the relationship between size, costs and outputs.

The simulations below use estimated coefficients from the fixed effects models in columns (2) and (5) of Table 5. We use all coefficients, regardless of their individual statistical significance.[29] Further, we hold constant the characteristics of the students served by comprehensive and themed schools. That is, students attend comprehensive and themed schools in the same patterns as they do in 2003. Using each school type's "own students," we then estimate costs for two different output levels; (a) current (2003) average outputs for each type, which do differ by type, and (b) aspirational outputs that are higher than the current output of schools – 90% graduation rate and 1000 combined SAT score – which we fix at the same level for both types.[30] The elasticity of costs with respect to size depends upon the quality of the outputs – say, the graduation rate – so that the optimal school size (i.e., the cost minimizing school size) will vary with the graduation rate. For example, a low graduation rate may mean costs are minimized at a large size while a high graduation rate might be cheaper to accomplish in a small school.[31]

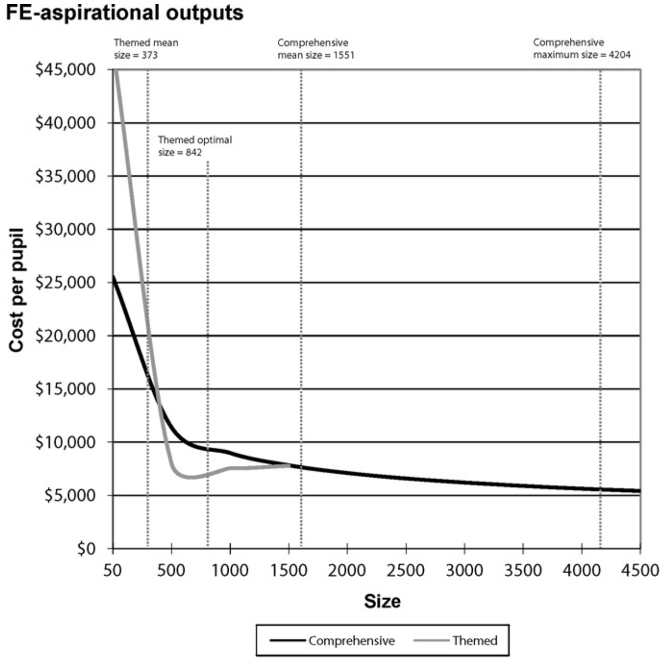
[27] We also estimate the equations in Table 5 using enrollment weights to counter possible heteroskedasticity. Results are qualitatively similar, although a few more coefficients in the fixed and random effects models show significance.

[28] We use the term "simulation" to describe the predictions of average costs based on alternative output levels for themed versus comprehensive schools.

[29] Joint *F*-tests on outputs and on student characteristics show significance at less than 7% in all cases.

[30] A second set of simulations relaxes the assumption of no change in student characteristics. Results were substantively similar.

[31] That is, the first derivatives are a function of the output levels.



Note: Simulations calculated using mean student characteristics by school mission type – comprehensive or themed - and aspirational outputs (90% graduation rate and 1000 combined SAT score).

Fig. 2. Simulations of high school costs for "own students" and aspirational outputs by type, 2003.

Fig. 1 graphs the cost curves of comprehensive and themed high schools based on the average mix of students and average output levels, respectively. To begin, notice that there are significant differences between the themed and comprehensive schools. Within the actual range of school enrollment in NYC, costs decline and then increase (a U-shaped relationship) for themed schools generally but for comprehensive schools costs decrease with enrollment suggesting cost-minimization outside the range of school sizes observed. Importantly, for the largest comprehensive schools, costs per pupil decrease very slowly with increases in enrollment so that any cost savings from expanding the largest schools will be minimal.

The key sizes which we highlight costs in Fig. 1 include the comprehensive maximum size in 2003 (4204 students), the comprehensive average size in 2003 (1551), the themed optimal size in 2003 (687), and the themed average size in 2003 (373).[32] At their respective average size, comprehensive schools cost less than themed schools – $7201 versus $8430 for a difference of $1229 per pupil – but both types of schools are currently smaller, on average, than their cost-minimizing size. At their maximum observed size of 4202 comprehensive high schools costs per pupil are $4996 or $2205 less than at their average size. At the themed optimal size of 687 students, costs per pupil are $7735 or $695 less than at their average size. Most notable, perhaps, is that comprehensive schools at either average or maximum observed sizes cost less than even optimally sized themed schools.

We next turn to our analysis of costs with our 'aspirational' output levels. Fig. 2 shows the relationship between costs and school size raising graduation rates and SAT scores to aspirational levels. The general shapes of the cost curves for comprehensive and themed schools are similar to those seen in Fig. 1 with two notable exceptions. First, the slope of the cost curve for comprehensive schools is flatter. Second, the slope of cost curve for themed schools is steeper with marginal cost savings greater up to the optimal size.

Moreover, at aspirational levels, the optimal size of themed schools is 842, which is 23% larger than the optimal size at average output levels (687) and 126% larger than the actual average size (373). Increasing outputs to aspirational levels will increase the costs per pupil in comprehensive schools, whereas for themed schools costs *decrease* when outputs are increased. Two possible explanations of this result for themed schools seem relevant. First, this may reflect the extent to which themed schools aim at producing outputs other than those captured here – 4-year graduation rates, SAT scores, etc. – such as increased college attendance or more creative students.[33] A second possible explanation lies in understanding the school enrollment decisions of students who have not graduated on time, that is, within the expected 4 years. These students may drop out of school entirely or continue to attend for as long as three

[32] See Appendix Table A.2 for the simulated cost estimates for the key enrollment points. Note also that we use the maximum observed size for the comprehensive schools, rather the optimal size which has roughly the same cost per pupil, but is out of sample.

[33] This might be particularly relevant for schools founded as alternatives to the traditional schools in the earliest wave of small schools, but less relevant for the newer schools that may have a more traditional academic focus.

additional years and, in fact, a significant portion of New York City's students graduate in 5, 6 or 7 years, rather than 4 years. Thus, if themed schools keep more students for a fifth, sixth or seventh year, then increasing the 4-year graduation rate may lower costs because it reflects a decrease in the 7-year graduation rate. Put differently, the explanation may lie in the differences in the length of time it takes students in themed and comprehensive schools to graduate. In themed schools, a lower percent of student graduate within 4 years, while a greater percent of students graduate in 5–7 years compared to comprehensive schools. Therefore, the more students that themed schools graduate in 4 years, the lower their costs.[34] In some sense, both of these explanations point to some unobserved additional output that we have not included. Although we have included more outputs than previous studies, it seems clear that future work with a richer set of outputs is warranted.

In summary, at key enrollment points, whether at the respective average output levels or aspirational levels, comprehensive schools cost less per pupil than themed schools. Further, some cost savings could be generated by expanding enrollment in comprehensive high schools, and increasing the size of themed schools. The optimal size of a themed school would still be a little more than half the size of the average comprehensive school. Perhaps most interesting, however, is that our results suggest that breaking up large high schools into smaller schools is unlikely to yield cost savings without concomitant changes in structure, policies and practices that characterize the themed schools. Smaller themed schools are cost-effective. Smaller comprehensive schools are not.

Clearly, policymakers are also likely to need more information than our study provides on the cost of capital investments related to creating new schools, be it building or leasing new physical space or reconfiguring existing space, as well as ongoing maintenance and transportation costs associated with the new schools. Nevertheless, information on operating costs, which account for the majority of the costs, is important.

7. Discussion: selection by students and teachers

Our discussion and analyses have thus far treated the characteristics of students as exogenous to the school type. It is, however, quite likely that there is at least some selection of students and, potentially, teachers into schools of different types and sizes. Understanding the nature and extent to which there is such selection is critical to drawing conclusions and recommendations for policy and practice. Further, while our analyses include variables aimed at controlling for differences in students, selection may yield differences in students and teachers that may shape our results. As an example, if better students choose to attend themed schools rather than comprehensive schools then our results may suffer from selection bias and the results would not be applicable to situations in which students are not able to choose what kind of school to attend.

To investigate the selection of students into type of schools, we obtained individual level data on a cohort of New York City 8th graders who attended a New York City high school in 9th grade. Our regressions indicate that there are, indeed, some differences between the students. As an example, probit regressions explaining whether a student attends a themed school indicate that the likelihood of attending a themed school is slightly higher for black and Hispanic students and also for those with low 8th grade test scores. A range of other variables is not statistically significant and the explanatory power of these variables is low (less than 5%). (Results available from the authors.) To control for these differences, however, we include the school level aggregates of all of these variables in the cost functions. Nevertheless, it seems clear that future work is needed to understanding the choices students make and the impacts of these choices on their educational outcomes.[35]

Selection by teachers is quite different from selection by students since we do not include the characteristics of teachers as regressors in the cost function. In general, cost functions include information on the prices of the inputs, rather than their quality, and in this context we might view the ability to attract and retain good teachers as a feature of cost-effectiveness that we are hoping to capture. That said, understanding this selection is critical to interpreting and applying our results. We perform two analyses to shed light on the question. First, we examine the factors determining the characteristics of the teachers in each school. Second, we investigate differences in the average salaries paid by themed and comprehensive schools.

To begin, we use data on four characteristics of teachers: the percentage licensed, percentage in school less than 2 years, percentage with more than 5 years of experience, and percentage with masters' degrees. We find that, on average, comprehensive schools have higher percentages of teachers with licenses, more than 5 years of experience and masters' degree, and lower percentages with less than 2 years in the school. However, regressions of these teacher characteristics on a set of school characteristics including school type by year suggest that school type *per se* is not an important factor. Coefficients on school type are statistically insignificant for licensure, experience and master's degree, although the school type is significant for the percentage with less than 2 years in some years, a likely reflection of the number of new themed schools. All of the independent variables used in these regressions are also included in our cost functions.

Notice, however, that the teachers may differ in other characteristics beyond the four for which we have measures. We explore this using data on teacher salaries as follows. Since all of the schools are in the same district, their teachers are paid according to same salary schedule which, as in other districts, determines teacher salaries based primarily on years of experience and education. Thus, dif-

[34] Five, six and seven-year graduation rates are not available at the school level for our panel.

[35] Note that it might also be possible that there are differences in the characteristics of students in schools of different sizes within the type category. This would, again, suggest controlling for these characteristics in the regressions as we have done.

ferences in average salaries across schools should be driven by differences in the years of experience and education of the teachers. We investigate the extent to which average teacher salaries differ in the themed and comprehensive schools controlling for the differences in the measures we observe, suggesting additional, unobserved differences in the teaching staff between the two types of schools. To be specific, we regress average teacher salaries on our teacher characteristics variables and a themed school dummy variable. (Results are not shown but are available from the author.) A positive (negative) coefficient on the themed school dummy would indicate that themed schools systematically employ teachers with unobserved characteristics that are better (worse) than the comprehensive schools. (Note that here better (worse) refers to factors that yield higher compensation.) The coefficient is, however, insignificantly different from zero, suggesting that unobserved compensated teacher characteristics are not systematically different between the two school types. Of course, this does not mean that they do not differ in uncompensated factors-say, energy, intelligence, charisma. Additional work examining the selection of teachers into schools is much needed to fully understand why and how these differences in costs emerge.

In summary, we find that observable differences across students predict the differences in teacher characteristics and in student choices across school types. Thus, controlling for these characteristics in the cost functions should substantially mitigate the possibility of bias in the estimates of the size coefficients, although more work in this area is clearly warranted.

8. Conclusions

A growing number of studies of the effects of high school size on various outputs indicate that smaller high schools result in better outputs. Only a few studies include costs along with outputs when studying the effect of school size. Yet cost-effectiveness in school reform is an essential analytic component, especially in large urban school districts with many competing ideas for reform and few ways to garner additional resources. Our analysis suggests that direct costs per pupil generally decline with size for all types of high schools. Importantly, moving to (small) themed schools from (large) comprehensive schools as they exist now will cost less per pupil as outputs increase. In particular, there could be cost savings for themed schools if they graduate more students in 4 years than 5–7 years.

Decreasing school size, however, is not enough—our findings show that small *comprehensive* high schools are expensive. If policymakers were to decrease school size and split comprehensive high schools into smaller schools, all they would accomplish is an increase in costs. The key to making small schools effective is to make them themed. While it is beyond the scope of this study to examine whether schools with particular themes are more or less costly in relations to size and outputs, it is important to better understand what it is about themed schools that does matter. Themed schools tend to have narrower course offerings in math, English and foreign languages with fewer Advanced Placement offerings as well. While this may be a limitation, it may also be an opportunity for more heterogeneity in classrooms within a more restricted range of offerings. Given that themed schools enroll students who are more difficult to educate, it is all the more impressive that themed small schools approach the costs of comprehensive schools at a certain point.

Our study finds that both size and mission matter, which is an important consideration for policymakers as they continue to seek ways to improve the educational outcomes of high school students. The pursuit of a singularly focused policy, such as creating small school without consideration of mission, will not produce the most cost-effective outcome. A more realistic approach would be to pursue a mix of schools in terms of both size and mission, understanding that the optimal size of both themed and comprehensive schools is larger than the average size of existing schools.

Acknowledgements

We thank the Carnegie Corporation for funding that supported this research. Sean Corcoran, William Duncombe, Henry Levin, Jennifer King Rice, Constancia Warren, Matthew Wiswall and seminar session participants at New York University's Institute for Education and Social Policy, the University of Kentucky's Martin School, and the American Education Finance Association's 2007 annual meetings contributed helpful suggestions. All findings and opinions are our own.

Definitions of small, medium, and large school sizes from the literature

See Table A.1.

Appendix A. School-based expenditure reports (SBER) functional categories (2003)

Total (Sum of I, II, III and IV)

I. Direct services (SUM of A, B, C, D, E, and F)
  A. Classroom instruction
    1. Teachers
    2. Educational para-professionals
    3. Other classroom staff
    4. Text books
    5. Librarians and library books
    6. Instructional supplies and equipment
    7. Professional development
    8. Curriculum development
    9. Contracted instruction
    10. Summer/eve school
  B. Instructional support services
    1. Counseling services
    2. Attendance and outreach
    3. Related services
    4. Drug prevent programs
    5. Referral/evaluation
    6. After school and student activities
    7. Parent involvement
  C. Leadership/supervision/support

1. Principals
2. Assistant principals
3. Supervisors
4. Secretaries and school aides
5. Supplies, materials, equipment and telephone

D. Ancillary support services
1. Food services
2. Transportation
3. School safety
4. Computer system support

E. Building services
1. Custodial services
2. Building maintenance
3. Leases
4. Energy

F. District support
1. Unscheduled sums/carry

II. District/superintendent costs (Sum A and B)

A. Instructional support and administration

B. Other districts and borough costs
1. Sabbaticals, leaves and termination pay
2. Additions to regular salary
3. Projected expenses

III. System wide costs (Sum of A and B)

A. Central instructional support

B. Central administration
1. Instruct offices
2. Operational offices
3. Central and chancellor's offices

IV. Other system wide obligations

A. Other system wide obligations
1. Debt service
2. Retiree health/welfare
3. Special commissioner for investigations
4. Projected expenses

Our dependent variable 'modified direct costs' is comprised of Direct Services (I), excluding Building Services (I.E.) and Transportation (I.D.2.)

Expenditure data from the New York City Department of Education (DOE) are not audited. The DOE assigns all public school expenditures to specific school locations, and only those schools recognized by the Department of Education are represented in the school-based expenditure reports. Some alternative programs and schools-within-schools, for example, are not reported as separate entities in the SBERs as they are not recognized as being distinct "expenditure locations" in the DOE's data systems. Spending on these programs is attributed to the "parent" school. Expenditures were attributed directly to schools when actual school allocation data and funds to support Direct Services to schools were available. Otherwise, indirect spending (such as Central Office) was allocated to schools on a per capita basis, unless spending was for particular populations (such as special education or Title I), in which case expenditures were allocated based on appropriate enrollment figures for each school. Expenditures were classified by function within the DOE's budget structure: Unit of Appropriation, Budget Code, Quick Code, Grant, Object code/Line Number. (Technical Appendix, School-Based Expenditure Reports, 2003)

Table A.1
Definitions of school size from the literature.

	NYC DOE; IESP small schools (Weinstein, Jacobowitz, Maguire, Saunders, & Fruchter, 2007)	Gates foundation (Gootman, 2005)	Stiefel et al. (2000)	NCES (2003)	Bickel et al. (2001)	Lee and Smith (1997)
Small	LE 500	LE 500	0–600	LT 300	220–550	LE 300, 301–600
Medium	NA	NA	600–1200; 1200–2000	300–599, 600–899	550–1800, 1800–3199	601–901, 901–1200
Large	NA	NA	GT 2000	GE 900	GT 3200	1201–1500, 1501–1800, 1801–2100, and GT 2100
Geographic area	NYC	NA	NYC	Nat'l.	TX	Nat'l. sample (NELS)

Table A.2
Simulations of high school cost, 2003.

Own students	Outputs	Cost per pupil	
		Comprehensive	Themed
Comprehensive maximum: 4204	Comprehensive, 2003	$4996	
Comprehensive average: 1551	Comprehensive, 2003	$7201	
Themed optimal: 687	Themed, 2003		$7735
Themed average: 373	Themed, 2003		$8430
Comprehensive maximum: 4204	Aspirational	$5544	
Comprehensive average: 1551	Aspirational	$7741	
Themed optimal: 842	Aspirational		$7504
Themed average: 373	Aspirational		$8743

Aspirational outputs include a 90% graduation rate and 1000 combined SAT score.

Simulated cost estimates at key enrollment points

See Table A.2.

References

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

Arenson, K. (2002). Gates to create 70 schools for disadvantaged. *New York Times*, March 19.

Bee, M., & Dolton, P. J. (1985). Costs and economies of scale in UK private schools. *Applied Economics*, *17*, 281–290.

Bickel, R., Howley, C., Williams, T., & Glascock, C. (2001). High school size, achievement equity, and cost: Robust interaction effects and tentative results. *Education Policy Analysis Archives*, *9*(40). <http://epaa.asu.edu/epaa/v9n40.html> (June 7, 2005).

Bloomfield, D. (2006). Come clean on small schools. *Education Week*, January 25.

Callan, S. J., & Santerre, R. E. (1990). The production characteristics of local public education: A multiple product and input analysis. *Southern Economic Journal*, *57*, 468–480.

Carnegie Corporation of New York (2006). *Education program*. Accessed November 20, 2006 from http://www.carnegie.org/sub/program/education.html.

Chabotar, K. J. (1989). Measuring the costs of magnet schools. *Economics of Education Review*, *8*, 169–183.

Cohn, E. (1968). Economies of scale in Iowa high school operations. *Journal of Human Resources*, *3*, 422–434.

Colegrave, A. D., & Giles, M. J. (2005). *School cost functions: A meta regression analysis*. Center for Labor Market Research, University of Western Australia Crawley, WA, paper.

Conant, J. B. (1959). *The American high school today*. New York, NY: McGraw-Hill.

Cotton, K. (1996). *School size, school climate, and student performance*. Portland, OR: NW Regional Laboratory. <www.nwrel.org/scpd/sirs/10/c020.html> (October 10, 2005).

Cotton, K. (2001). *New smaller learning communities: Findings from recent literature*. Portland, OR: NW Regional Laboratory.

Foreman-Peck, J., & Foreman-Peck, L. (2006). Should schools be smaller? The size-performance relationship for welsh schools. *Economics of Education Review*, *25*(2), 157–171.

Fowler, W. J., Jr. (1992). What do we know about school size?-What should we know? In *Paper presented at the annual meeting of the American Educational Research Association* San Francisco, CA,

Fowler, W. J., Jr., & Walberg, H. J. (1991). School size, characteristics, and outcomes. *Educational Evaluation and Policy Analysis*, *13*(2), 189–202.

Gewertz, C. (2006). Chicago's small schools seen gains, but not on tests. *Education Week*, August 9.

Gootman, E. (2005). Gates donates $28.5 million for support of small schools. *The New York Times*, February 5.

Iatarola, P., Schwartz, A. E., Stiefel, L., & Chellman, C. (in press). Small schools, large districts: Small school reform and New York City's students. *Teachers College Record*, *110*(9).

Kumar, R. C. (1983). Economies of scale in school operation: Evidence from Canada. *Applied Economics*, *15*, 323–340.

Kuziemko, I. (2006). Using shocks to school enrollment to estimate the effect of school size on student achievement. *Economics of Education Review*, *25*, 63–75.

Lee, V. E., & Smith, J. B. (1997). High school size: Which works best, and for whom? *Educational Evaluation and Policy Analysis*, *19*(3), 205–227.

National Center for Education Statistics (2003). *The condition of education*, NCES 2003-067. Washington, DC: US Government Printing Office. <http://nces.ed.gov/pubsearch/pubsearchinfo.asp?pubid=2003067> (April 1, 2005).

Oakland Unified School District (2000). *New small autonomous schools district policy*. Revised Draft, May 15 <www.ousd.k12.ca.us/news/New_Schools/Revised_NSA_%20Policy.htm> (November 9, 2000).

Page, L., Layzer, C., Schimmenti, J., Bernstein, L., & Horst, L. (2002). *National evaluation of smaller learning communities: Literature review*. Cambridge, MA: Abt Associates, Inc. February <www.abtassociates.com/reports/SMALLER.pdf> (September 23, 2005).

Ravitch, D. (2006). Downsize high schools? Not too far. *Washington Post*, November 6, 7(B).

Raywid, M. A. (1996). *Taking stock: the movement to create mini-schools, schools-within-schools, and separate small schools*. New York: ERIC Clearinghouse on Urban Education, Teachers College, Columbia University. (No. ED396-045)

Raywid, M. A. (1999). *Current literature on small schools. ERIC Digest*. Charleston, WV: ERIC Clearinghouse on Rural Education and Small Schools. (No. ED4250-49).

Ready, D. D., Lee, V. E., & Welner, K. G. (2004). Educational equity and school structure: School size, overcrowding, and schools-within-schools. *Teachers College Record*, *106*(October (10)), 1989–2014.

Riew, J. (1966). Economies of scale in high school operation. *Review of Economics and Statistics*, *48*, 280–287.

Riew, J. (1986). Scale economies, capacity utilization and school cost: A comparative analysis of secondary and elementary schools. *Journal of Education Finance*, *11*, 433–440.

Schneider, B., Wyse, A. E., & Keesler, V. (2007). Is small really better? Testing some assumptions about high school size. In L. Tom, & F. M. Hess (Eds.), *Brookings Papers on Education Policy 2006/2007*. Washington, DC: The Brookings Institution.

Schwartz, A. E., & Stiefel, L. (2006). Is there a nativity gap? The achievement of New York City elementary and middle school immigrant students. *Education Finance and Policy*, *1*(1), 17–49.

Stiefel, L., Berne, R., Iatarola, P., & Fruchter, N. (2000). High school size: Effects on budgets and performance in New York City. *Educational Evaluation and Policy Analysis*, *22*, 27–39.

Viadero, D. (Mary 23, 2006). Smaller not necessarily better, school-size study concludes. *Education Week*, accessed May 31, 2006 from www.edweek.org/ew/articles/2006/05/23/38samall_web.h25.htm.

Watt, P. A. (1980). Economies of scale in schools: Some evidence from the private sector. *Applied Economics*, *12*, 235–242.

Weinstein, M. G., Jacobowitz, R., Maguire, C., Saunders, T. B., & Fruchter, N. (2007). *Stability in student and teacher characteristics in the first ten years: A study of small high schools in New York city*. New York University Institute for Education and Social Policy.