ELSEVIER

Using shocks to school enrollment to estimate the effect of school size on student achievement

Ilyana Kuziemko*

Department of Economics, Harvard University, Cambridge, MA 02138, USA

Abstract

Previous studies of the connection between school enrollment size and student achievement use cross-sectional econometric models and thus do not account for unobserved heterogeneity across schools. To address this concern, I utilize school-level panel data, and generate first-differences estimates of the effect of school size on achievement. Moreover, to account for the possibility that trends in both achievement and enrollment size are jointly determined, I exploit shocks to enrollment provided by school openings, closings, and mergers in a two-stage-least-squares estimation. The results suggest that smaller schools increase both math scores and attendance rates and that the benefit of smaller schools outweigh the cost.

1. Introduction

As the number of enrolled students at the typical primary or secondary school in the US has grown for decades, small schools have become an increasingly popular topic among politicians, educators and parents. Encouraging the creation of smaller student bodies became a popular policy during President Clinton's second term, culminating in the Small Schools Initiative proposed in Clinton's 2000 State of the Union Address. Organizations such as the Annenberg Foundation, the Carnegie Corporation, the Pew Charitable Trusts, and the Bill and Melinda Gates Foundation (which announced a $240 million investment to promote smaller schools) have all stated their support for shrinking the size of America's public schools, arguing that smaller schools will foster closer ties between teachers, students and parents and will improve student achievement. And a recent survey by the nonprofit opinion research firm Public Agenda found that 66% of parents and 80% of teachers felt that schools are currently too big.

Parents and educators may be reacting to the way that school consolidation has transformed the American educational landscape during the 20th century. In 1920, 271,000 schools served students in the US, while in 1990 only 83,000 remained. As the population grew robustly throughout this period, the average enrollment-per-school increased by a factor of six during this period.[1]

*Tel.: +1 617 493 2222; fax: +1 617 495 7730.
E-mail address: kuziemko@nber.org.

---

[1]See *The Biennial Survey of Education (1920–1922)* (Department of the Interior, Bureau of Education, 1922) and *The Digest of Education Statistics (1991)*.

But does school size really matter to educational outcomes? Economists have had almost nothing to say on this question. Almost all of the current work on the subject is from the education literature, and is poorly identified. The lack of economic research on the subject (compared, for instance, to the wealth of work linking *class* size and student achievement) is surprising as the field has left a major social phenomenon unexplored. In the large set of papers in the education literature on the question of school size, more studies have found negative than positive effect of size on achievement, but no consensus has emerged. In a review of the literature on economies of size in education, Andrews, Duncombe and Yinger (2002) lament the lack of credible research on school consolidation: "Since program evaluation research on school consolidation is limited, it is time for researchers on both sides of the debate to make good evaluation research on consolidation a high priority."

Much of the current confusion in the literature may be due to the empirical weaknesses that the existing papers share. Each employs a cross-sectional specification, which will yield inconsistent estimates if the variation in enrollment size across schools is not exogenous. For example, suppose that the largest schools are in poor, urban areas, where many other unobserved factors (such as a lack of parental involvement or positive role models, or the presence of underachieving peer groups) may also negatively influence student achievement. Then cross-sectional estimates of the effect of enrollment on student achievement are likely to be biased negatively. On the other hand, if large schools tend to be in wealthy suburbs and small schools tend to be in poor rural areas, the bias would be positive. In general, it is difficult, a priori, to sign the bias in a cross-sectional estimate, and there would be little stability in the estimates across samples and specifications.

The aim of this paper is to correct the bias that has made the results of past papers difficult to interpret. To isolate the effect of school size from that of any unobserved variables, I employ school-level panel data for elementary schools in Indiana between 1989 and 1998. Previous cross-sectional studies exploit variation across schools at a single moment in time; I exploit variation across time within each school in a first-differences estimation. While this method addresses the omitted-variables problem that plagues the existing literature, its estimates of the school-size effect still may not be consistent. If *trends* in both enrollment and student achievement were jointly determined—which would be the case if, say, schools that had improving test scores also attracted a growing number of students—then conventional first-differences or fixed-effects estimates would be biased. To address this concern, I use school mergers, openings, and closings as instruments for changes in enrollment. These events serve as shocks

to the number of students attending a school; I use these deviations-from-trend to measure the effect of enrollment size on student achievement.

The paper is organized as follows. Section 2 reviews the existing literature on the school size question. Section 3 describes the data. Section 4 explains the estimation strategy. Section 5 discusses the results. Section 6 compares the costs and benefits of smaller schools. Section 7 concludes and offers directions for future research.

2. Literature review

The literature on the topic of school size follows two broad strands. Some papers develop theories as to why there should be an effect of school size on student performance. Others try to quantify the size of such an effect.

2.1. Why school size might affect student achievement

The justifications for school consolidation depend chiefly on economies-of-scale arguments. Conant (1959) was the first to point to the potential benefits of large, "comprehensive" schools. Conant and his followers highlighted the variety of classes that can be offered when demand from a large group of students is pooled; the increased specialization of teachers afforded by finer divisions of labor; the decreased average costs per student when 50, not five, students can use the same piece of expensive laboratory equipment and when food and other non-durables can be purchased in bulk.

Supporters of large schools point to potential psychological benefits as well. Smith and DeYoung (1988) make the following arguments: first, small schools are more likely to draw only from a limited geographical area, and are thus more likely to be demographically homogenous. If exposure to students of different backgrounds is considered a function of education, then school consolidation might be justified on grounds of diversity. Second, interaction with a limited number of teachers may be stifling to students. If a student has the same teacher over the course of several years, the teacher's expectation of the student may be set after the first year of such contact, precluding student development. In a larger school, where teachers are specialized to teach only one grade-level, students get a fresh start each year. Third, limited enrollment may hinder a student's social development. In a larger school, with many social groups or "cliques," a student may have a better chance of finding one such group in which he feels comfortable.

The arguments against school consolidation counter many of the above claims. Several studies have undermined the notion that large schools offer more diverse

curricula. Barker and Gump (1964) first questioned this benefit of large schools, but their evidence was largely anecdotal. Using a 1984–1985 dataset of New York State public high schools, Monk (1987) found that increases in high school size beyond 1500 students effect no change in the absolute number of classes offered. Moreover, he found no correlation between size and the quality of classes offered: calculus and other Advanced Placement classes were equally available in small schools. Elsworth (1998) finds similar results looking at the senior-year course offerings of Australian secondary schools.

Another concern of small-school proponents is the lack of extra-curricular opportunities in large schools. Though large schools may have a greater absolute number of sports teams and clubs, they usually have a smaller rate of extra-curricular involvement per student. This is especially true for the more competitive activities such as varsity sports: there are only 12 spots on the basketball squad, regardless of school size, but in a large school, 500 students compete for them, while only 100 do so in a smaller school. Coladarci and Cobb (1996) find that extra-curricular involvement is significantly higher in smaller schools. They use the National Education Longitudinal Study of 1988 and find a large, negative and statistically significant effect of school-size on extra-curricular participation, though their controls for socio-economic status are minimal.

Finally, small-school advocates address the psychological arguments made by the large-school camp. Strang (1987) calls attention to the alienating effects of large, bureaucratic schools. The specialization larger schools provide, they claim, comes at a price: a student now has five or six teachers, none of whom knows the student very well. Indeed, much of the recent enthusiasm for smaller schools was generated after the shootings at Columbine High School, whose large size was seen as hampering teachers' ability to recognize the psychological troubles of the two gunmen. Walberg and Walberg (1994) point out the stronger connection small schools foster between student and community. Finally, small schools may engender old-fashioned but effective learning styles in the classroom. Palincsar and Brown (1986) document the benefits of the "one-room school" practices of mixed-age grouping, peer tutoring, and reciprocal learning in which students teach one another.

2.2. Measuring the school-size effect

Almost all of the existing studies that measure the school-size effect follow roughly the same econometric strategy. These studies use a cross-sectional specification and do not account for unobserved heterogeneity among schools. While they differ in the exact dependent variable used (most employ standardized test scores, graduation rates or college-entrance rates), most make use of the same explanatory variables: the number of enrolled students and some basic controls for the socio-economic status of students' families. Fowler and Walberg (1991), Lee and Smith (1995) and Deller and Rudnicki (1993) are illustrative of the current literature. All of these papers find a negative effect of school size on student achievement.

Other studies have found the opposite effect. Sander (1993) finds a positive effect of school size on ACT scores[2] and Barnett, Glass, Snowden, and Stringer (2002) find that large schools are more cost-effective. Bradley and Taylor (1998) deserve special attention because they not only examine cross-sectional regressions but also use a first-differenced specification, finding a positive effect of enrollment on exam scores in both cases. However, trends in enrollment and test scores might also be jointly determined (e.g., improving districts could be located in growing neighborhoods), which would bias first-differenced results.

Finally, Lamdin (1995) finds no effect of school size on student performance using data from the Baltimore school district in 1990. He argues that because unobserved variables that tend to bias the coefficient on school size would vary less within a district than across districts, limiting the sample to schools within the same district minimizes such bias. However, if better schools attract a larger enrollment, then the effect of school size could be positively biased. Parents might transfer their children to better schools (using either a formal school-choice program or "Tiebout-style" residential choice), making them larger. While the costless mobility assumption of the Tiebout model is unlikely to hold, as long as some parents are sensitive to school quality in their location decision, better schools could, all else equal, attract a larger student body. The bottom line is that while using observations from a single district may improve estimates, without plausibly exogenous variation in school size it is difficult to interpret the coefficients in a cross-sectional regression.

3. The data

From the Indiana Department of Education, I obtained school-level information on achievement test scores and attendance. Each fall since 1988, the state has administered the Indiana Statewide Test for Educational Progress (ISTEP) to every public school student in the 3rd and 6th grades. The test-score data contain the average score on each section of the exam for every

---

[2]ACT scores are not an ideal dependent variable because the ACT is not a required exam. A school's average score is a function not only of the academic achievement of the student body, but also of the sample of students choosing to take the exam. For this reason, it is difficult to interpret Sander's results.

Table 1
Summary statistics

Variable	Mean	Standard deviation	Minimum value	Maximum value
Average daily attendance rate	0.9599	0.0107	0.8653	0.994
Average ISTEP math score	62.94	8.828	6.5	97.62
Average ISTEP language score	62.68	7.671	8.0	89.92
School enrollment	418.3	169.7	36	1487
White share of enrollment	0.862	0.224	0	1
Black share of enrollment	0.107	0.208	0	1
Hispanic share of enrollment	0.020	0.0509	0	0.804
Asian share of enrollment	0.0065	0.0139	0	0.467
Share of enrollment receiving federally subsidized school lunch	0.256	0.194	0	1

public school in the state. The attendance data from the state Department of Education give the average daily attendance rate for each public school in the state.

I merged this data with the Public School Universe data from the National Center for Educational Statistics (1991, 2000, 2002). This data set provides enrollment and student-body demographics for every public school in the nation. The relevant summary statistics from both data sets appear in Table 1.

From the Public School Universe data, I was able to identify abrupt changes in school enrollment. Such changes fall into one of two categories. First, in some cases, several schools merged to form a larger school, or one school split into several new schools. Second, in the other cases, a school opened (closed) which reduced (increased) the enrollment of other schools in the district. Fig. 1 illustrates the repercussions of a typical shock. In 1995, Orchard View Elementary opened, drawing primary students from the rest of the Middlebury Community School District. A sharp decline in the enrollment of the older schools (Jefferson, York, and Middlebury Elementary Schools) resulted.

Over 100 schools in the Indiana public school system experienced such shocks to their enrollments between the 1988–1989 and 1998–1999 school years. As most of these changes took place at the elementary school level, I focus on these younger students. I use as dependent variables third grade test scores in the mathematical and language portions of the exam as well as average daily attendance rates in elementary schools. As noted by Lamdin (1996), attendance may best be thought of as an educational input, not an output, and thus does not belong on the left-hand side of a regression. However, as one of the key claims of the small-school camp is that large schools increase absenteeism[3] (both by increasing commute times and by creating overwhelming and alienating environments), I will consider the direct effect on attendance.
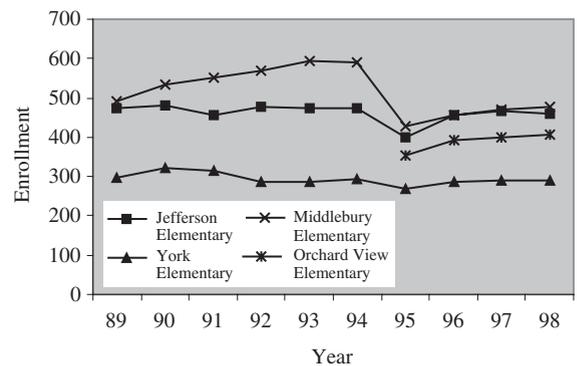
Fig. 1. The effect of a typical shock on elementary school enrollment.

4. Empirical strategy

The literature review revealed that previous studies of school size use a cross-sectional econometric specification. The drawback to this strategy is that it does not account for unobserved heterogeneity across schools. If an unobserved variable were positively related to achievement and negatively (positively) related to school size, then regressing student achievement on school size would yield an estimate of the coefficient on school size that is negatively (positively) biased.

I have two strategies that I will use to address this problem. The first strategy is conventional first-differences estimation: this regression implicitly compares the same school at different points in time instead of comparing different schools to each other. Specifically, I regress the change in an achievement variable between consecutive years on the change in enrollment (as well as the change in the demographic composition of each school and a set of year dummy variables). The validity of this method depends on whether changes in enrollment are exogenous. If a school's enrollment changes randomly over time then this specification should provide unbiased estimates of the effect of school size; if it responds to the quality of the school (perhaps

_____

[3]See, for example, Carnie (2002).

Table 2
First-stage estimation for attendance 2SLS regressions

Variable	Changes over one year	Changes over two years	Changes over three years
Change in enrollment due to a shock	0.9948***	1.012***	0.9815***
	(0.0270)	(0.0533)	(0.0518)
Change in white share	−0.00926	−0.169**	−0.185*
	(0.0497)	(0.0683)	(0.0999)
Change in black share	0.0777	0.1273*	0.100
	(0.0552)	(0.0748)	(0.105)
Change in Asian share	0.461***	0.436**	0.552***
	(0.131)	(0.173)	(0.208)
Change in Hispanic share	−0.0712	−0.212**	−0.285**
	(0.0871)	(0.104)	(0.128)
Change in free-lunch share	−0.159***	−0.249***	0.00030***
	(0.0196)	(0.0256)	(0.0000234)
Observations	9802	8587	7483
R-squared	0.135	0.0657	0.0856

parents move to a certain area because they know that the schools are improving) then this approach will not provide unbiased estimates.

My second and preferred strategy is to use the "shocks" of mergers and school openings and closings as instruments for school enrollment changes in a two-stage-least-squares (2SLS) regression model. In the first stage, I regress enrollment changes on a variable that equals the enrollment change of a school if that change resulted from a "shock" and zero otherwise, and on other exogenous variables (the year dummies and the change in demographic composition). In the second stage, I use this estimate of enrollment changes in the regression of achievement indicators on enrollment, demographic changes, and year dummies. The panel nature of this specification addresses possible unobserved-variables bias; the use of instrumental variables addresses possible simultaneity between changes in enrollment and changes in school quality.[4]

Because the effects of enrollment size may take a few years to materialize, I lag the dependent variables one, two, and three years. That is, I estimate the following equation:

$$Y_{i,j+k} - Y_{ij} = \alpha(N_{i,j+1} - N_{ij}) + \beta(D_{i,j+1} - D_{ij}) + \gamma(W) + \varepsilon_{ij}, \quad (1)$$

[4]Note that mergers generate especially good instruments because they do not suffer from "composition effects." When a school closes and its students are added to a neighboring school, that neighboring school is flooded with a new group of students. Thus, comparing the test scores in the neighboring school before and after the other school closed is not ideal because the pool of students taking the exam has significantly changed over time. Conversely, comparing the (weighted) average of the test scores of two schools before they merge with the test scores of the new, merged school allows assessment of the same students over time.

where i indexes the school, j indexes the year, Y is the school average of student achievement variables (ISTEP scores or attendance rates), N is enrollment, D is a vector consisting of the share of the student body that is white, black, Asian, and Hispanic, and the share that receives free federal school lunches, W is a vector of year dummies, and ε is the error term. I estimate the equation separately when k equals one, two, and three. Thus, I am identifying the effect of an enrollment change on student achievement the year after the change, as well as two and three years later.

Finally, it is not clear which unit of measurement is the most appropriate for measuring the change in student achievement over time and the change in enrollment over time. I use the absolute change in the dependent variable (as the test scores and attendance rates are defined as percentages or percentiles originally) and the percent change in enrollment. Though I do not report them here, the results reported in the next section are robust to using the absolute change in enrollment and the percent change in test scores and attendance rates.

4.1. A closer look at the enrollment instrument

Two necessary conditions for a variable to be a valid instrument are: (1) that it is correlated with the endogenous variable; and (2) that it is uncorrelated with the error term.

The first condition is easily demonstrated. Intuitively, it is obvious that a shock to school enrollment caused by a merger or school closing will affect a school's population by exactly the size of the shock itself. That is, if a school were to gain 50 students because a neighboring school closed, the expected change in the population of that school the following year would be 50 students, all else equal. Table 2 shows the first-stage

Table 3
Enrollment shocks to Indiana elementary schools, 1988–1999

Variable	Negative enrollment shock	Positive enrollment shock
Number of schools experiencing shocks	58	39
Median absolute change in school enrollment due to shock	−140	+110
Median percent change in school enrollment due to shock	−22.7%	+29.0%
Median growth rate of school the year before a shock	+2.75%	−1.52%
Median enrollment of school the year before a shock	606.5	355
Median growth rate of district elementary student population the year of a shock	+2.44%	−3.21%

regressions.[5] The effect of an enrollment shock on total enrollment is one-for-one and highly significant, just as we would expect.

As is usually the case, demonstrating that the second condition holds is more difficult. My identification strategy is that of *regression discontinuity*: I focus on the shocks to school enrollment and compare student performance before and after these shocks.[6] Of course, the enrollment shocks to Indiana public schools are not random: they are responses to underlying population trends in each school district. However, the shocks do provide sharp discontinuities in the enrollment trend of a school. A school with growing enrollment or a school in a district with a growing enrollment is the most likely candidate for a shock—such as a new school opening in the district—that decreases its enrollment; similarly, a school with a shrinking enrollment is more likely to experience a positive shock to its enrollment. Table 3 shows characteristics of schools that have experienced enrollment shocks and characteristics of the districts in which they are located. Note that schools experiencing positive enrollment shocks have a median enrollment of 339, have shrinking student enrollments, and are located in districts with shrinking enrollments; schools that experience negative shocks have a median enrollment of 596, have growing student enrollments, and are located in districts with growing enrollments. Thus, most shocks not only provide a discontinuity in the enrollment trends of the district and the school, but are contrary to the trend itself.

That enrollment shocks are determined by community population growth may seem troubling in that community population trends might also be related to school

quality, which would bias the IV results. However, while enrollment shocks are hardly random, the decision of exactly when to undertake them essentially is. As Table 3 shows, enrollment shocks are responses to long-term trends, and the decision to respond to these trends in year t as opposed to $t-1$ or $t+1$ is due to enrollment reaching a certain threshold over or under which the district decides it must act. The specific year enrollment reaches that threshold is in large part random, and it is this variation that the instrument exploits.

5. Results

For the sake of comparison, Table 4 displays the results from a pooled, cross-sectional regression using math scores as the dependent variable. Clearly, the results are not stable across specifications: as controls are added, both the sign and magnitude of the coefficient on enrollment changes. Because of the endogeneity of the enrollment variable, the effect of adding or omitting relevant covariates will depend on their correlation with both the achievement and enrollment variables—correlations whose signs we may not know a priori.

For this reason, I focus on methods that make use of the panel nature of the data. The conventional first-differenced estimation results and the 2SLS results with attendance rates as the dependent variable appear in Table 5. Parallel results for math and language test scores appear in Tables 6 and 7, respectively. In nearly all of the regressions, the coefficient on the enrollment variable is negative and significant. The results for the 2SLS specification using changes over two and three years are highly statistically significant.

In the first-differences estimations, we see a consistent negative effect of enrollment on achievement, statistically significant in most specifications. In the 2SLS regressions, there is little significant effect of enrollment on achievement in the following year. However, two and three years later we see negative and statistically significant effects on attendance and math scores, and a negative (but not statistically significant) effect on language scores.

---

[5] The first-stage regressions reported in Table 2 are from the attendance estimations in Table 4, though the first-stage results from the math and language estimations are essentially identical.

[6] van der Klaauw (2002) provides an excellent discussion of regression discontinuity as well as an application to measuring the effects of college financial aid. Angrist and Lavy (1999) and Hoxby (2000) have used similar methodologies to examine the connection between class-size and student achievement.

Table 4
Cross-sectional estimates of the effect of enrollment on math scores

Variable	Coefficient (standard error)		
Enrollment	0.001106**	−0.000508	−0.000776*
	(0.000467)		(0.000407)
Free-lunch share of students		−22.827**	−21.27**
		(0.349)	(0.488)
Black share of students			−0.437
			(0.440)
Hispanic share of students			−11.92**
			(1.21)
Asian share of students			56.74**
			(5.05)
Observations	12093	12093	12093
R-squared	0.0005	0.2612	0.2737

Notes: Coefficients significant at the 0.1 level are denoted by *, and those at the 0.05 level by **.

Table 5
The effect of enrollment on attendance rates

Variable	Specification					
	First differences			Two-stage least squares		
Change in enrollment	−0.0020**	−0.00109	−0.000944	0.00254*	−0.00306*	−0.0042**
	(0.00071)	(0.000713)	(0.000692)	(0.00150)	(0.00188)	(0.00184)
Change in white share of enrollment	0.00139	−0.001709	−0.001625	0.00157	−0.00173	−0.0016
	(0.00257)	(0.003932)	(0.002703)	(0.00258)	(0.00393)	(0.00275)
Change in black share of enrollment	−0.001646	−0.00411	−0.00319	−0.00188	−0.00398	−0.00295
	(0.00369)	(0.00434)	(0.00342)	(0.00370)	(0.00433)	(0.00349)
Change in Hispanic share of enrollment	−0.00744	−0.0148**	−0.001814	−0.00680	−0.0150**	−0.00234
	(0.00577)	(0.00743)	(0.00608)	(0.00602)	(0.00728)	(0.00634)
Change in Asian share of enrollment	0.00204	−0.00626	−0.01016	−0.000143	−0.00526	−0.00845
	(0.00623)	(0.00789)	(0.007559)	(0.00606)	(0.00813)	(0.00799)
Change in free-lunch share of enrollment	−0.0059**	−0.00300*	−0.00222	−0.0052**	−0.0033**	−0.00269
	(0.001653)	(0.00167)	(0.001817)	(0.00168)	(0.00169)	(0.00184)
Change in dependant variable over how many years?	One	Two	Three	One	Two	Three
Observations	9802	8580	7386	9802	8580	7386
R-squared	0.0258	0.0302	0.0381	0.0192	0.0292	0.0358

Notes: Coefficients significant at the 0.1 level are denoted by *, and those significant at the 0.05 level are denoted by **. All regressions include year fixed-effects. Schools resulting from mergers are weighted in proportion to the enrollments of the original school the year prior to the merger. Standard errors (in parentheses) are corrected for clustering on the new school created after a merger. The regressions in columns four, five and six were re-run with the full set of instruments (in order to perform Hausman specification tests). The coefficients (and standard errors) on the enrollment variables are, following the ordering in the table: 0.00291 (0.00183), −0.00349 (0.00206), and −0.00420 (0.00176).

Thus, both strategies suggest a negative effect of school size on achievement. The main difference between the results of the two approaches is that in the regressions using dependent variables two and three years in the future, the coefficients on enrollment in the conventional first-differences regressions are consistently smaller in absolute value than those on the enrollment coefficients in the 2SLS regressions. If parents relocate in order to enroll their children in better schools, then schools with improving test scores might have growing admissions; similarly, those with falling test scores might have shrinking admissions. Thus, simultaneity between trends in test scores and enrollment might positively bias the coefficient on enrollment in the first-differences regressions.

The 2SLS estimates indicate that school size has a meaningful effect on student achievement. The point-estimates suggest that doubling enrollment leads to a

Table 6
The effect of enrollment on math scores

Variable	Specification					
	First differences			Two-stage least squares		
Change in enrollment	−1.447**	−1.18*	−1.2724	−1.20	−3.841**	−4.123*
	(0.6449)	(0.7257)	(0.8828)	(2.09)	(1.427)	(2.250)
Change in white share of enrollment	−0.4774	−4.46**	−1.386	−0.460	−4.601**	−1.4201
	(1.640)	(1.969)	(2.601)	(1.63)	(1.963)	(2.6025)
Change in black share of enrollment	−6.263**	−9.133**	−4.417	−6.27**	−8.983**	−4.2051
	(2.281)	(2.458)	(2.939)	(2.302)	(2.383)	(2.9439)
Change in Hispanic share of enrollment	−5.995	−8.913*	0.9670	−5.97	−9.328*	0.42724
	(4.814)	(5.100)	(5.269)	(4.79)	(5.230)	(5.6401)
Change in Asian share of enrollment	−2.071	0.6167	−0.0761	−2.182	1.948	1.3855
	(5.648)	(5.001)	(5.013)	(5.748)	(5.013)	(5.2522)
Change in free-lunch share of enrollment	−1.077	−2.506**	−3.759**	−1.039	−2.876**	−4.196**
	(1.385)	(1.253)	(1.464)	(1.407)	(1.257)	(1.5282)
Change in dependant variable over how many years?	One	Two	Three	One	Two	Three
Observations	10733	9429	8328	10733	9429	8328
R-squared	0.1764	0.2647	0.2816	0.1763	0.2633	0.2802

Notes: Coefficients significant at the 0.1 level are denoted by *, and those significant at the 0.05 level are denoted by **. All regressions include year fixed-effects. Schools resulting from mergers are weighted in proportion to the enrollments of the original school the year prior to the merger. Standard errors (listed in parentheses) are corrected for clustering on the new school created after a merger. The regressions in columns four, five and six were re-run with the full set of instruments (in order to perform Hausman specification tests). The coefficients (and standard errors) on the enrollment variables are, following the ordering in the table: −0.570 (1.67), −4.99 (1.74), and −5.97 (2.58).

Table 7
The effect of enrollment on language scores

Variable	Specification					
	First differences			Two-stage least squares		
Change in enrollment	−1.04**	−0.8302	−0.9825	−0.2734	−1.187	−1.656
	(0.5214)	(0.5382)	(0.67103)	(1.115)	(1.029)	(1.765)
Change in white share of enrollment	5.936**	0.5590	3.2078	5.99**	0.5404	3.1998
	(2.078)	(1.957)	(2.3108)	(2.088)	(1.957)	(2.3156)
Change in black share of enrollment	0.6764	−4.689**	−1.189	0.6321	−4.669**	−1.1395
	(2.207)	(2.393)	(2.7448)	(2.214)	(2.375)	(2.7172)
Change in Hispanic share of enrollment	−0.4093	−5.0878	4.2503	−0.3369	−5.144	4.1228
	(4.355)	(3.675)	(4.0775)	(4.338)	(3.693)	(4.1593)
Change in Asian share of enrollment	5.223	10.982**	−5.4097	4.880	11.16**	−5.0643
	(5.490)	(4.969)	(4.577)	(5.457)	(5.01)	(4.652)
Change in free-lunch share of enrollment	0.6540	−2.530**	−3.433**	0.7727	−2.579**	−3.537**
	(1.096)	(1.0673)	(1.180)	(1.094)	(1.0722)	(1.209)
Change in dependant variable over how many years?	One	Two	Three	One	Two	Three
Observations	10733	9429	8328	10733	9429	8328
R-squared	0.1729	0.1722	0.2594	0.1727	0.1722	0.2592

Notes: Coefficients significant at the 0.1 level are denoted by *, and those significant at the 0.05 level are denoted by **. All regressions include year fixed-effects. Schools resulting from mergers are weighted in proportion to the enrollments of the original school the year prior to the merger. Standard errors (listed in parentheses) are corrected for clustering on the new school created after a merger. The regressions in columns four, five and six were re-run with the full set of instruments (in order to perform Hausman specification tests). The coefficients (and standard errors) on the enrollment variables are, following the ordering in the table: −0.291 (1.36), −1.456 (1.291), and −2.314 (2.077).

4.1%-point decrease in math scores and a 0.4%-point decrease in attendance three years later. Alternatively, they suggest that a one standard-deviation increase in enrollment is associated with a 0.15 standard-deviation decrease in math scores and a 0.17 standard-deviation decrease in attendance rates three years later. I will address the economic relevance of these results in the next section, but I note here that my estimates compare favorably with even the most generous estimate of the effect of reducing class size.[7]

5.1. Discussion and extensions

A possible objection to the results in this section is that educational inputs such as class size are not included in the regressions. Unfortunately, only the first three years of data contain the total number of teachers per school. Note, however, that not including class size will only bias the 2SLS results if changes in class size are correlated with changes in school size when schools experience enrollment shocks. Using the three years of data that are available, I find that changes in enrollment do not explain changes in student-to-teacher ratios in years that schools experience a shock.

Another objection to the 2SLS specification is that the first-stage estimation assumes that changes in demographic composition are exogenous. To test this assumption, I perform a Hausman specification test, comparing the results of the original 2SLS regressions to those of 2SLS regressions that used the demographic composition changes caused by enrollment shocks as instruments for demographic changes (as well as retaining the original instrument for enrollment changes). The coefficients on the enrollment variables in the new regressions are not statistically different from (and, in fact, are strikingly similar to) those of the original 2SLS model, which indicates that the assumption of the exogeneity of the demographic changes is benign.[8] Moreover, the results in these tables are robust to including no demographic controls, only the subsidized lunch control, or only the race controls.

One explanation for the negative coefficient on the enrollment variable in the 2SLS regressions is that the disruption generated by enrollment shocks—and not the increase in enrollment itself—is driving the drop in achievement indicators. Most of the negative changes in enrollment are triggered by the opening of new schools, while most of the positive changes are triggered by two or more schools merging into one. If mergers are inherently more difficult than new school openings for teachers, principals, and students, then the *nature* of the enrollment shocks, and not the change in student enrollment that they generate, could be the underlying reason for the negative coefficient on the enrollment variable.

However, closer investigation of the results undermines this explanation. First, this objection does not address the negative effect of school size on achievement found in the conventional first-differenced results, in which most of the variation in school size comes from changes in community population, not from enrollment shocks. Second, the results from the 2SLS regressions do not support this explanation. If the enrollment effect worked only through the confusion generated in the wake of a merger, then this effect should be most prevalent in the first year and weaken in subsequent years. Instead, the enrollment effect seems to take at least a year to materialize. Even the attendance results do not show negative effects of enrollment in the first year, which would be expected if the underlying reason for lower attendance was that students simply travel further to their school when the pool of students it serves grows. Indeed, for math scores and attendance rates, the negative effect of enrollment as measured by the 2SLS regressions tends to grow in absolute value each year after an enrollment change, suggesting that the longer students attend larger (smaller) schools, the more their achievement indicators fall (rise).

In a final extension, I experiment with non-linear models of the effect of enrollment on student achievement. I estimate 2SLS regressions using the usual dependent variables and the absolute change in enrollment, this value squared, and the usual controls as explanatory variables.[9] As Table 8 shows, these regressions did not provide any conclusive results. However, each dependent variable gives rather consistent signs for the coefficients on the enrollment terms: negative on the linear term and positive on the quadratic term. When I solve to find the "worst" enrollment, I get solutions between 540 and 7000. Taken literally, these results suggest that, after a certain threshold, increasing school size improves outcomes. However, as this threshold is far above the enrollment levels for most of the schools in the sample, the vast majority of schools would be predicted to improve if their enrollments were *decreased*, not increased.

---

[7]See Krueger (1999), which represents the upper bound on the effect of class size on achievement. He estimates that a 50% increase in class size is associated with a 0.20 standard deviation fall in math scores. Using values from Table 6, I find that a 50% increase in school size is associated with a 0.23 standard-deviation fall in math scores.

[8]I include the coefficients and standard errors from these regressions in the footnotes to Tables 4–6.

---

[9]Note that these regressions use the absolute change in enrollment and not the percentage change. This is done to ease the interpretation of the coefficients.

Table 8
2SLS regressions with quadratic terms for change in enrollment

Variable name	Change in achievement variables over two years			Change in achievement variables over three years		
	Attendance	Math	Language	Attendance	Math	Language
Change in enrollment	−0.0000364**	−0.00813*	−0.02943**	−7.93e-6	−0.00707	−0.01715
	(0.0000157)	(0.00427)	(0.01179)	(0.0000228)	(0.00518)	(0.01765)
Change in enrollment squared	2.84e-8**	−4.10e-6	0.000027**	−3.21e-9	5.12e-7	0.000013
	(1.31e-8)	(2.85e-6)	(0.00001)	(2.30e-8)	(3.16e-6)	(0.0000153)
Change in percent White	−0.001426	−4.594**	0.8584	−0.002055	−1.346	3.3310
	(0.003933)	(1.843)	(1.926)	(0.0029484)	(2.611)	(2.3631)
Change in percent Black	−0.004543	−7.839**	−5.271**	−0.002452	−3.726	−1.3903
	(0.004357)	(2.328)	(2.312)	(0.003744)	(3.015)	(2.7231)
Change in percent Hispanic	−0.011868	−9.2408*	−1.9604	−0.004093	1.2562	5.5883
	(0.007618)	(5.484)	(4.307)	(0.008372)	(5.798)	(4.627)
Change in percent Asian	−0.004732	−0.20590	11.448**	−0.002347	0.3713	−4.876
	(0.008347)	(5.258)	(5.120)	(0.007891)	(5.175)	(4.679)
Change in percent free lunch	−0.003506**	−2.5898**	−2.7883**	−0.002547	−4.388**	−3.6165**
	(0.001704)	(1.3428)	(1.0847)	(0.001852)	(1.548)	(1.2102)
Enrollment value for which $\frac{\partial(Test\ Scores)}{\partial(Enrollment)} = 0$	640.84	Derivative positive for all values of *Enrollment*	539.1	Derivative positive for all values of *Enrollment*	6905.27	659.6

6. Is decreasing school size worth the cost?

While a full cost-benefit analysis of school size is beyond the scope of this paper, I use the approach of Krueger (2002) to do a rough, back-of-the-envelope calculation of the costs and benefits. Note that throughout the following analysis I look only at the individual benefits to a representative student in terms of future income. If student achievement exhibits positive externalities, then this estimate will fall short of the total benefit to society.

Consider a 50% decrease in school size. Using the results in Table 5, we would expect math scores to rise by about 2% points, or one quarter of a standard deviation, over two years. Currie and Thomas (1999) find that a one standard deviation increase in math test scores is associated with a 7.6% increase in future earnings, while Neal and Johnson (1996) find that a one standard deviation increase in the Armed Forces Qualification Test is associated with a 20% increase in future earnings. As a conservative estimate, I equate a one standard deviation increase in math scores with an 8% increase in future earnings, so that our representative students would receive a 2% increase in earnings from the proposed policy change.[10]

The average personal income in the US of a full-time, year-round worker in 2001 was about $44,848.[11] Assuming that there is zero real income growth (an implausible assumption, but one that will bias the results towards finding a smaller benefit), that the discount rate is 4% (roughly equal to the return on a long-term inflation-indexed US government bond), and that individuals work from the age of 18–65, then the total present discounted value of the *additional* income a student will receive due to the policy change is $12,744.[12]

---

[10]The Currie and Thomas (1999) result is especially relevant for my calculations as they use scores from exams that respondents took at the age of seven—roughly the age of the

(*footnote continued*)

students in my sample. The Neal and Johnson result applies only if I assume that the test-score gains students made in elementary school persist in high school.

[11]From the author's calculations in 2001 dollars using Current Population Survey data.

[12]Assume that our representative third grader (aged 8) enters the labor force at age 18 and retires at age 65 and has a constant yearly income Y of $ 44,848 if school size remains at current levels. As we assume that the treatment effect of smaller schools, b, is a 2% increase in income, the present discounted value of the increase in future income streams equals

$$b|Y \sum_{t=18}^{65} \frac{1}{(1+r)^{t-10}} = (0.02)|\$44,848 \sum_{t=18}^{65} \frac{1}{(1.04)^{t-10}} \approx \$12,744$$

To calculate the costs of smaller schools, I consider separately the changes in the operating costs and changes in the capital costs resulting from new school construction. Using the upper-bound estimates from Taylor and Bradley (2000), I calculate that a 50% decrease in enrollment size leads to a 20% increase in per-pupil operating costs. Using the national average of per-pupil spending in 2001 of $7524,[13] I estimate that the per-pupil increase in operating costs associated with six years of instruction (the length of elementary education) in a school that is 50% smaller is equal to $8,204.[14]

Of course, new schools will have to be built to enact such a policy change. There were 69,697 public elementary schools in the US in 2000, serving 33,709,000 students.[15] Assuming the plan is enacted immediately, about 34,849 new elementary schools would have to be built. Abramson (2000) finds that the average cost of public school construction in 2000 was about $8,006,000 (in 2001 dollars). The average age of current schools is about 40 years (which is obviously a lower bound on the average life of each school).[16] Hence, the per-pupil annual increase in costs from school construction would be roughly $207, which leads to a total of $1242 over the six-years of elementary instruction. Hence, the per-student net present value of the policy change (benefits minus costs) is about $3298, in 2001 dollars.

Note that throughout this section, I have generally used values and assumptions that bias the result in the direction of finding a smaller benefit and a larger cost. For instance, my regressions estimate the benefit of attending smaller schools for *two* years; the effect of attending a smaller school for all six years of elementary education would most likely be higher. Yet I calculate costs assuming students spend all six years of their elementary education in costlier, smaller schools. While this exercise is meant to be exploratory in nature and has made strong assumptions about both the linearity and the stability of cost and benefit functions across different studies and data sets, it does suggest that reduction in

school size may be a cost-effective way to increase student achievement. If the benefits of smaller schools can be achieved through relatively cheap reforms such as "schools-within-schools," then educators may be able to reap the benefits of smaller schools without losing the savings generated by economies of scale.[17]

7. Conclusion

The results of the first-differences and 2SLS regressions indicate that, as a first approximation, reducing school size increases student achievement as measured by average daily attendance rates and standardized math scores. The negative effects of enrollment increases do not appear to be the result of any short-run disruption associated with the enrollment shocks themselves. In fact, the effect of changes in enrollment does not appear the year after a shock, but two and three years later, suggesting that the longer students spend in larger schools, the greater the decline in their achievement indicators. The results from the quadratic specification suggest that the effect of school size is greater at small values of enrollment and is probably non-linear.

Further research would benefit from a larger and more detailed data set. A larger data set would allow for more precise estimates of the non-linearities in the enrollment effect, as well as provide more enrollment shocks at the middle-school and high-school levels so that the effect of enrollment on older students could be quantified. Additionally, a larger data set would allow separation of urban and suburban schools without reducing the number of enrollment shocks to the point where a viable instrument could not be generated. Thus, any difference in the effect of school size in urban versus suburban areas could be identified. Finally, a dataset with more information on school inputs such as class size could determine more conclusively if shocks to school size are systematically accompanied by changes in per-pupil resources—a question this paper only partially addresses.

A more detailed data set would also broaden the scope of outcomes that could be investigated. While I only had access to average test scores, future studies could focus on other moments of the distribution of scores. As small schools may be better at identifying students who are "falling through the cracks," school size may have an especially large effect on the left-hand tail. Moreover, if small schools are less able to track, enrollment could affect the variation in student performance if tracking helps high-ability students and hurts

---

[13]See *Digest of Education Statistics, 2002*, Table 166. As this figure includes both operating and capital costs, it will overstate the increase in operating costs.

[14]Let c be the original per-pupil cost of education and let g be the percentage increase due to the policy change. Then, the additional cost equals

$$\sum_{t=1}^{6} \frac{cg}{(1+r)^{t-1}} = \sum_{t=1}^{6} \frac{\$7524(0.2)}{(1.04)^{t-1}} \approx \$8,204$$

[15]See *Digest of Education Statistics, 2002*, Tables 40 and 87.
[16]See *Condition of America's Public School Facilities: 1999*.

[17]See Dewees (1999) for a review of the school-within-a-school literature.

low-ability students, as in Argys, Rees, and Brewer (1996).[18]

This paper suggests that policy-makers' recent interest in school size is not misplaced, and that there seems to be a causal connection between school size and the performance of elementary school students. While setting school size unreasonably low may not be financially feasible, administrators should consider the benefits to students of smaller enrollments when determining the size of their schools.

Acknowledgements

References

Abramson, P. (2000). School planning and management construction report, 2000. *School Planning and Management*, *39*(2), 17–34.

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*(3), 245–262.

Angrist, J. D., & Lavy, V. (1999). Using Maimonides' Rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics*, *114*(2), 533–575.

Argys, L., Rees, D., & Brewer, D. (1996). Detracking America's schools: equity at zero cost? *Journal of Policy Analysis and Management*, *15*(4), 623–645.

Barker, R., & Gump, P. (1964). *Big school, small school*. Palo Alto, CA, Stanford: University Press.

Barnett, R., Glass, J. C., Snowden, R., & Stringer, K. (2002). Size, performance, and effectiveness: cost-constrained measures of best-practice performance and secondary-school size. *Education Economics*, *10*(3), 291–311.

Bradley, S., & Taylor, J. (1998). The effect of school size on exam performance in secondary schools. *Oxford Bulletin of Economics and Statistics*, *60*(3), 291–324.

Carnie, F. (2002). Small is beautiful: lessons from America. *Education Revolution*, *36*, 36–39.

Coladarci, T., & Cobb, C. (1996). Extracurricular participation, school size, and achievement and self-esteem among high school students: a national look. *Journal of Research in Rural Education*, *12*(2), 92–103.

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

Currie, J., & Thomas, D. (1999). Early test scores, socio-economic status and future outcomes. *NBER Working Paper # 6943*.

Deller, S., & Rudnicki, E. (1993). Production efficiency in elementary education: the case of Maine public schools. *Economics of Education Review*, *12*(1), 45–57.

Department of the Interior, Bureau of Education (1922). *Biennial survery of education*, 1920–1922. Washington, DC.

Dewees, S. (1999). The school-within-a-school model. *ERIC Digest*.

Elsworth, G. (1998). School size and diversity in the senior secondary curriculum: a generalizable relationship? *Australian Journal of Education*, *42*(2), 183–203.

Figlio, D., & Page, M. (2000). School choice and the distributional effects of ability tracking: does separation increase equality? *NBER Working Paper # 8055*.

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

Hoxby, C. (2000). The effects of class size on student achievement: new evidence from population variation. *The Quarterly Journal of Economics*, *115*(4), 1239–1285.

Krueger, A. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*, *114*(2), 497–532.

Krueger, A. (2002). Economic considerations and class size. *NBER Working Paper # 8875*.

Lamdin, D. (1995). Testing for the effect of school size on student achievement within a school district. *Education Economics*, *3*(1), 33–42.

Lamdin, D. (1996). Evidence of student attendance as an independent variable in education production functions. *Journal of Educational Research*, *89*(3), 155–162.

Lee, V., & Smith, J. (1995). Effects of high school restructuring and size on early gains in achievement and engagement. *Sociology of Education*, *68*(4), 241–270.

Monk, D. (1987). Secondary school size and curriculum comprehensiveness. *Economics of Education Review*, *6*(2), 137–150.

National Center for Education Statistics. (1991). *Digest of educational statistics*. Washington, DC.

National Center for Education Statistics. (2000). *Condition of America's public school facilities: 1999*. Washington, DC.

National Center for Education Statistics. (2002). *Digest of educational statistics*. Washington, DC.

Neal, D., & Johnson, W. (1996). The role of pre-market factors in black–white wage differentials. *Journal of Political Economy*, *104*, 869–895.

Palincsar, A., & Brown, A. (1986). Interactive teaching to promote independent learning from text. *Reading Teacher*, *39*(8), 771–777.

Sander, W. (1993). Expenditures and student achievement in Illinois: new evidence. *Journal of Public Economics*, *52*(3), 403–416.

Smith, D., & DeYoung, A. (1988). Big school vs. small school: conceptual, empirical, and political perspectives on the re-emerging debate. *Journal of Rural & Small Schools*, *2*(2), 2–11.

---

[18]The majority of papers have found that tracking increases the variation of student achievement; however, see Figlio and Page (2000) for a good discussion of the identification issues in this literature.

Strang, D. (1987). The administrative transformation of American education: school district consolidation, 1938–1980. *Administrative Science Quarterly, 3*(3), 352–366.

Taylor, J., & Bradley, S. (2000). Resource utilization and economies of scale in secondary schools. *Bulletin of Economic Research, 52*(2), 123–150.

van der Klaauw, W. (2002). Estimating the effect of financial aid offers on college enrollment: a regression-discontinuity approach. *International Economic Review, 43*(4), 1249–1287.

Walberg, H., & Walberg, H. (1994). Losing local control. *Educational Researcher, 23*(5), 19–26.