Educational Evaluation and Policy Analysis
Summer 1999, Vol. 21, No. 2, pp. 193–213

The Behavioral Effects of Variations in Class Size: The Case of Math Teachers

Julian R. Betts
University of California-San Diego

Jamie L. Shkolnik
National Opinion Research Center

This paper tests whether variations in class size cause teachers to alter their teaching methods. Examination of 2,170 math classes suggests that when class size is reduced, teachers do not spend more time on new material, nor do they finish more of the assigned textbook. Rather, teachers shift time away from group instruction and towards individual instruction. Time spent on student discipline and routine administration declines modestly, while time spent on review rises. Even large reductions in class size shift teachers' time allocation by only a few percentage points. Teachers react more strongly to class size changes when teaching below-average students.

1. Introduction

California and other states have recently implemented programs to reduce class size in public schools for one or more grades. These recent announcements in fact merely continue a longstanding trend in the United States toward smaller classes.[1] In spite of this trend, research evidence on what reduction in class size can achieve is mixed. Hanushek (1996) reviews 277 estimates of the impact of the teacher-pupil ratio and finds that only 15 report positive and statistically significant effects on student outcomes, compared to 13 that report significant negative outcomes.[2] Direct experimental evidence on the impact of class size reduction (CSR) has emerged from the STAR experiment in Tennessee. See, for example, Finn and Achilles (1990), Moesteller (1995), and Krueger (1997). The STAR experiment suggested that reduced class sizes increased test scores for elementary school students, but most of the gains came in the first year in which students experienced the CSR. The lack of a similarly sized impact on learning in later years has puzzled researchers, while raising questions about unobserved aspects of student placement into the control and experimental groups in the initial year of the study.

Examination of the long-term effects of class size on students' earnings much later in life has produced a similarly ambiguous message. Betts (1996) reviews the literature, and finds that even the most optimistic studies imply a very small internal rate of return to spending on CSR. In addition, studies that measure school resources at the school actually attended rather than relying on imputation based on a person's state of birth tend to produce smaller estimated effects of smaller pupil-teacher ratios on students' subsequent earnings.

Adding to the uncertainty about CSR is the lack of information about the exact mechanism(s) through which CSR might improve student outcomes. The issue of how teachers respond to smaller class sizes lies at the heart of the question. As Murnane (1991) argues, "at its core, education consists of the interaction of human beings. In particular, unless additional funding results in a change in someone's behavior, no real changes in outcomes will occur" (p. 462).

This paper attempts to get inside the black box of CSR to understand how teachers do or do not change their teaching methods when faced with a change in class size. We use a detailed panel data set on classes taught by math teachers in Grades 7–12, sampled from a representative national sample of American public schools. We explore how a variation in class size affects the teachers' allocation of time between group and individual instruction. We exploit a second set of questions answered by teachers in which they indicate how they allocate class time in percentage terms among time spent on new material, review, discipline, routine administrative matters, and testing. Finally, we explore whether teachers with smaller classes succeed in finishing a greater percentage of the assigned math textbook by the end of the school year. The answers to these questions promise to provide useful new insights into how CSR might, or might not, change the teaching behavior of mathematics teachers in American middle schools and high schools.

The next section discusses these issues from a theoretical perspective, and touches upon some of the relevant findings from earlier work. Section 3 describes our data set, Section 4 presents the details, and Section 5 concludes.

2. The Theory of Time Allocation and Class Size Reduction

As mentioned above, we will consider three distinct types of teacher behavior. The first involves an analysis of how teachers divide their time between group instruction and individualized instruction. The second type of behavior is measured by a set of variables indicating the teacher's estimate of how much time he or she devotes to various tasks inside the classroom, each of which may involve individual or group instruction. Finally, we have what might best be considered an outcome variable that reflects teachers' choices indirectly: the percentage of the textbook covered.

Consider first the question of how teachers allocate time between group instruction time (which we will denote as G) and an average of individual instructional time per student (I). In this case, if a teacher has T minutes to teach each week, and if there are N students in the class, then the time budget constraint is

$$(1) \qquad T = G + I \times N$$

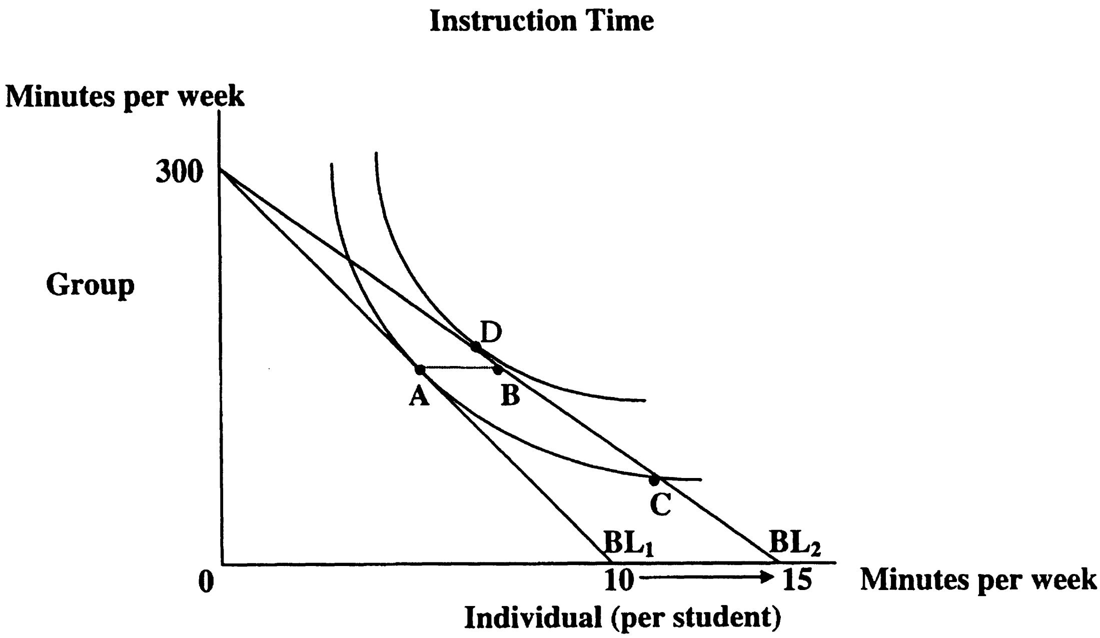
Suppose that the teacher's goal is to maximize average student learning. (We briefly discuss below what might happen if, instead, the teacher's goal is to improve the achievement of the less advanced students in the class.) Assume that student achievement S is given by

$$(2) \qquad S = f(G, I)$$

where the signs of the first and second derivatives of the production function f are given by $f_G > 0, f_I > 0, f_{GG} < 0, f_{II}$ and $f_{GI} \geq 0$. Figure 1 shows the possible results of a reduction in class size. The straight line is the time budget constraint, with T set to 300 minutes in our example. It shows all combinations of group and individual instruction time from which a teacher can choose. Initially, class size is 30, so that if a teacher devoted all his or her time to individual instruction, $I = 300/30 = 10$. The convex curve tangent to the budget line at point A is an isoquant[3] showing combinations of I and G that produce equal levels of student achievement. As is standard in economic theory, we assume that this curve is convex.[4] Since it takes more time to produce better student achievement, isoquants that are further to the northeast of the origin represent higher levels of achievement. Point A on the figure shows the optimal combination of G and I: Any other combination of G and I on the budget line would produce a lower level of student achievement. Mathematically, the solution at point A is defined by the equality of the ratio of the marginal products of time spent on G and I with the relative "price" of the two types of teaching. Since one minute of individual instruction to each of N students in the class necessitates a drop in group instruction of N minutes, the "price" of one minute of individual instruction per student is N minutes of group instruction. At the optimum a minute of individual instruction should be N times as effective as a minute of group instruction:

$$(3) \qquad \frac{MP_I}{MP_G} = \frac{f_I}{f_G} = N$$

The figure also shows one of many possible outcomes when class size is reduced from $N = 30$ to $N = 20$. In the new equilibrium, the teacher moves to point D, increasing both G and I. But it is possible that if the isoquants were shaped differently, the optimal solution would be at point B, so that the teacher would leave group instruction time un-



Note. With a decrease in class size, the budget line rotates out from BL_1 to BL_2.

FIGURE 1. *Teacher's maximization problem with a choice between group and individual instruction.*

changed, but would increase individual instruction per student.[5]

The model underlying the figure assumes that teachers aim to maximize average student rates of learning. How could this analysis be squared with the frequent finding in the literature that CSR has no, or very little, effect on mean student achievement? One possibility is that after a reduction in class size, the teacher moves from point A to point C, leaving average student achievement unchanged. Presumably, a teacher might increase individual per student instruction, I, to a large extent, as shown, if he or she placed a greater value on helping the weakest students in the class. Brown and Saks (1987) provide direct evidence that teachers do prefer to grant more individual instruction time to students at the bottom of the class. Shkolnik and Betts (1998) find that when class size is reduced, teachers increase I by more than they should if their goal were truly to maximize average student achievement.

A second set of variables that we will model does not distinguish between group instruction and individual instruction, but instead simply asks the teachers what percentage of class time they allocate to each of five activities (time spent on new material, review, discipline, routine administrative matters, and testing). In the empirical section, we will begin our analysis of these variables by labeling discipline and routine as 'non-instructional time' and the remaining three time uses as 'instructional time.' Unfortunately, the survey does not ask teachers to divide each of these activities into group and individual instruction. But suppose that these five activities are denoted X_j, $j = 1,\ldots,5$, and that a more fully specified version of the test score equation in (2) can be written

$$S = f(G, I, X_1, \ldots, X_5) \quad (4)$$

Recognizing from the above analysis that G and I are functions of class size N, we could write a partially reduced form of this model as:

$$S = f(N, X_1, \ldots, X_5) \quad (5)$$

The way in which the teacher apportions total time spent on each of the five tasks will be governed by a slightly different time budget function from before:

$$T = X_1 + X_2 + X_3 + X_4 + X_5 \quad (6)$$

The important difference between this optimization problem and the earlier problem of allocating T between I and G is that the relative price, that is, the tradeoff between time spent on one activity and another, is always equal to one, not N. So when class size changes, it does not rotate the budget line as in Figure 1. Now, the optimal allocation of time will be such that the marginal product of one minute spent on any activity j just equals the marginal product of one minute spent on any of the remaining activities i:

$$(7) \qquad \frac{\partial f}{\partial X_1} = \frac{\partial f}{\partial X_2} = \cdots = \frac{\partial f}{\partial X_5}$$

One might think that this solution implies that a change in class size will not affect the optimal allocation of time to the five activities. However, a change in class size is almost sure to cause the optimizing teacher to reallocate his or her time. The reason is that an increase in class size is likely to affect the marginal products of each of the five inputs, listed above, by different proportions. To make the example concrete, suppose that when class size increases from 20 to 30, the teacher finds than an extra minute spent testing students becomes more valuable, because he or she now has less direct evidence available on how students are coping with the coursework. At the same time, the value of time spent on new material might fall, since the students are no longer learning as quickly. The only way to reallocate teaching time so as to maximize student learning is to increase time spent on testing while reducing time spent on new material. Under the assumption of diminishing returns ($\partial^2 f / \partial X_j^2 < 0$ for all j), this reallocation of time will continue until the marginal products of all the time inputs are re-equalized. Seen in this light, it becomes almost certain that a teacher whose goal is to maximize student achievement will respond to a change in class size by changing his or her allocation of time among the five types of time use.

Finally, we will analyze how variations in class size affect the percentage of the textbook finished during the course, as reported by the teacher. This dependent variable differs from the others in our analysis in that it is not a measure of time allocation. Rather, it should be viewed as an outcome variable, reflecting the decisions that the teacher makes about how to allocate time between G and I, and between time spent on new material, review, testing, and so on.

3. Data

The data are taken from the Longitudinal Study of American Youth (LSAY), a national study that followed students from approximately 100 nationally representative middle schools and high schools during the period from fall 1987 to spring 1992. The LSAY includes surveys completed by principals, teachers, students, and parents, and provides detailed information on student and teacher background characteristics, student test scores, and student and teacher behavior in the classroom.

Every spring, for each class in which one or occasionally more LSAY student participants were enrolled, math teachers filled out a survey describing the math class that they taught the previous fall. We created a data set with one observation per classroom by discarding duplicate observations, after matching on a large series of variables provided by the teacher about the classroom. We obtained a data set of roughly 2,170 observations, each describing a class taught by a math teacher. The variables include school information such as overall enrollment and demographics. This information was obtained from a survey of principals, a survey of the teachers regarding their work experience and education as of 1987, and an annual series of information forms detailing the nature of each class that included one or more LSAY students. In many cases, teachers appear more than once in our data because they taught more than one class to a student included in the LSAY.

4. Results

Table 1 displays the means and standard deviations of the variables used in the analysis. In order to obtain a sample that was relatively uniform across variables, we restricted the data in this table to observations that had information for at least one of the dependent variables and for all of the explanatory variables. The table shows that mean class size for the classes in the sample was 23.7 with a fairly large standard deviation of 6.9. The average class lasted 301.9 minutes per week, again with a large standard deviation.

The nine dependent variables in our analysis are shown at the bottom of Table 1. On average, teachers finished 79.7% of the textbook, devoted 140.3 minutes to group instruction, and devoted 8.5 minutes of individual instruction to each student. Of total class time, teachers on average devoted 86.6% to instructional activities: 30.7% was

TABLE 1
Means and Standard Deviations for All Variables

	M	*SD*	No. of observations
Class size	23.65	6.88	2,173
Total minutes/week	301.88	85.23	2,173
Class ability level	3.40	1.17	2,173
Lunch assistance at school (%)	13.69	15.11	2,173
Black students at school (%)	15.11	20.66	2,173
Hispanic students at school (%)	3.01	6.26	2,173
School size	1,136.00	521.29	2,173
Suburban	0.48	0.50	2,173
Rural	0.28	0.45	2,173
Female students in class (%)	49.05	12.72	2,173
Teacher has MA degree	0.56	0.50	2,173
Teacher experience (years)	15.33	8.35	2,173
7th grade	0.09	0.29	2,173
8th grade	0.10	0.30	2,173
9th grade	0.12	0.32	2,173
11th grade	0.27	0.44	2,173
12th grade	0.10	0.29	2,173
Text covered (%)	79.71	14.67	2,109
Group instruction (min/week)	140.32	53.45	2,173
Individual instruction/student (min/week)	8.45	12.57	2,173
Class time on instructional activities (%)	86.61	9.08	2,159
Class time on discipline (%)	6.14	6.88	2,158
Class time on review (%)	30.66	13.01	2,164
Class time on new material (%)	40.19	14.36	2,166
Class time on routine (%)	7.27	4.50	2,166
Class time on testing (%)	15.77	6.74	2,161

devoted to review, 40.2% was devoted to new material, and 15.8% was devoted to testing. Non-instructional time consisted of the 7.3% of total class time devoted to routine administration and the 6.1% devoted to discipline. All of these variables displayed large variations relative to the mean.

We now model our measures of teacher time allocation as a function of class size and other variables. We want to isolate the impact of variations in class size on teacher time allocation, while controlling for confounding variables that might be correlated with both teachers' time allocation and class size. Serious misspecifications could result unless we control for teachers' characteristics, the types of classes they teach, the types of students they teach, and overall characteristics of the school and the student body. For instance, in a school or classroom with large numbers of less advanced students, teachers may feel compelled to spend relatively large amounts of time on individual remediation and on review as opposed to new material. Since in this and other data sets, classes for less advanced students tend to be smaller, our estimates of how class size affects individual instruction and review might be biased downward if we did not control for the achievement level of students in the class. Accordingly, we add a class ability variable (on a scale of 1–5, with 5 indicating the most able classes) based on each teacher's estimate of the average achievement level of students in their class.[6, 7] Betts and Shkolnik (in press) and Shkolnik and Betts (1998) both find that controlling for class ability leads to much larger and more significant estimates of the impact of CSR on students' rate of learning. In addition, we add "0,1 variables" (dummy variables) for 20 of the 21 types of math classes taken by students in the sample, to control for systematic differences in how teachers teach courses ranging from consumer math to calculus.

We also control for overall traits of the student body at the school—the percentage of students receiving full lunch assistance; the percentage who are Black or Hispanic; school size measured in total enrollment; and dummy variables for suburban (equal to one if the school is in a suburban area, zero otherwise) and rural schools (equal to one if the school is in a rural area, zero otherwise). Urban schools, the comparison group, is the omitted category. We also add dummy variables for each grade level in the study (7^{th}–12^{th}) with Grade 10 as the omitted category or the comparison group.

It remains quite possible that teachers vary systematically in how they allocate their time in the classroom, and that this might be spuriously correlated with the typical class sizes they teach. We therefore control for whether the teacher holds a Master's degree by including a dummy variable for teacher education. The teacher education variable equals one if the teacher holds a Master's degree, zero otherwise. We employ a quadratic in years of teaching experience (that is, we include both experience and experience squared) based on several findings in the test-score literature which suggest that teachers become more effective during their first few years, and then level off in performance. It stands to reason that this non-linear effect of teachers' experience on student learning could reflect non-linear changes in how teachers run their classrooms as they gain experience.)

We begin in Table 2 with Ordinary Least Squares (OLS) regressions. Column 1 shows the percentage of the text covered by the teacher. A reasonable working hypothesis here is that teachers with smaller classes finish a greater proportion of their textbooks. Surprisingly, this does not appear to be the case. Furthermore, teachers with more class time minutes per week do not finish significantly greater percentages of their textbooks. (This finding is somewhat less surprising, in that teachers who know well in advance that they can count on relatively large amounts of class time at their school might adopt lengthier textbooks.) The single most important variable in the regression is the (1–5) class ability level. A top-level class is predicted to finish an additional 8.8% of the assigned textbook, compared to a bottom-level class. Variations in class size do little to compensate for this difference. We also find evidence that teachers with Master's degrees finish about an extra 1.3% of their textbooks, on average.

Again, we stress that the percentage of text covered represents an outcome that is a combined result of teacher and student behavior and school resources. If it is true that class size has little impact on how far teachers proceed through the typical textbook, could this reflect a small or non-existent impact of class size on how teachers teach? Columns 2 and 3 answer this question more directly by examining how variations in class size affect teachers' allocation of time to group instruction and individual (per student) instruction.[8] Here we find statistically significant evidence that in larger classes, teachers substitute group instruction for individual instruction. An increase in class size from 20 to 30 students is predicted to increase time spent on group instruction by 4.8 minutes, while reducing individual instruction per student by 7.4 minutes. The measured impact on I is larger because it is composed of two changes—the overall 4.8 minute drop in total $I \times N$, and the quite mechanical drop in individual time that results from spreading a given number of minutes between a greater number of students.

These represent some of the most impressive shifts in teachers' methods that we find resulting from variations in class size. Shkolnik and Betts (1998), using a similar extract from the LSAY, found that group instruction time has larger effects on student learning than does individual instruction, and that teachers overreact to CSR by reallocating too much time to individual instruction. In that paper, we also found that after controlling for class ability, CSR was predicted to increase students' rate of learning. However, the effects would grow considerably greater if teachers held G constant instead of decreasing G after a reduction in class size.

The final column in the table models the percentage of class time teachers devoted to instructional activities (as opposed to non-instructional activities, which consist of time spent on discipline and time spent on daily routine). The results indicate that an increase of 10 in class size reduces instructional time by about 2%.

Table 2A shows the relevant coefficients when the models in Table 2 are repeated with the addition of a squared class size term. Our goal is to test for varying responses as class size continues to grow. For group and individual instruction, we find significant weakening of the effects of class size as class size increases. Increases in class size are predicted to increase group instruction for class sizes of up to 38.2, while individual instruction per student, I, is predicted to fall for class sizes of up to 35.3[9]. A more interesting finding is found in the model of the percentage of the text covered: Both

TABLE 2
Ordinary Least Squares Regressions

	Text covered (%)	Group instruction (min/week)	Individual instruction/student (min/week)	Time on instructional activities (%)
Constant	75.086	37.638	13.780	84.143
	(27.568)	(4.467)	(6.66)	(51.693)
Class size	0.0692	0.4811	–0.7441	–0.1986
	(1.387)	(3.138)	(–19.720)	(–6.690)
Total minutes/week	0.0023	0.2972	0.0350	–0.0008
	(0.653)	(26.923)	(12.892)	(–0.364)
Class ability level	2.1987	–0.2546	0.0312	1.5483
	(6.223)	(–0.233)	(0.116)	(7.335)
Lunch assistance at school (%)	–0.0850	0.2482	–0.0511	0.0706
	(–2.380)	(2.293)	(–1.918)	(3.357)
Black students at school (%)	–0.0350	–0.0249	0.0191	–0.0575
	(–1.685)	(–0.395)	(1.230)	(–4.700)
Hispanic students at school (%)	0.1460	–0.6150	0.1266	–0.2214
	(2.309)	(–3.173)	(2.654)	(–5.845)
School size	–0.0010	0.0092	–0.0002	0.0015
	(–1.386)	(4.100)	(–0.434)	(3.380)
Suburban	0.1167	2.1629	0.4069	–0.3074
	(0.136)	(0.817)	(0.624)	(–0.599)
Rural	–1.7037	–6.9542	1.8857	–0.7128
	(–1.754)	(–2.314)	(2.550)	(–1.220)
Female students in class (%)	–0.0094	–0.0103	0.0082	0.0334
	(–0.382)	(–0.136)	(0.438)	(2.262)
Teacher has MA degree	1.3198	2.2345	0.1659	1.4805
	(1.914)	(1.051)	(0.317)	(3.600)
Teacher experience (years)	–0.1784	–1.2647	0.0449	–0.2285
	(–1.364)	(–3.150)	(0.455)	(–2.945)
Teacher experience squared	0.0040	0.0325	–0.0029	0.0064
	(1.010)	(2.678)	(–0.963)	(2.710)
7th grade	1.5974	–0.3354	2.0277	6.2690
	(0.520)	(–0.035)	(0.859)	(3.388)
8th grade	0.8554	–10.264	1.2457	4.4460
	(0.475)	(–1.878)	(0.926)	(4.214)
9th grade	–0.0028	2.8095	0.2752	–0.3150
	(–0.003)	(0.841)	(0.335)	(–0.487)
11th grade	–0.2627	0.1357	1.3973	–0.6320
	(–0.316)	(0.052)	(2.191)	(–1.261)
12th grade	–0.8967	–0.2314	0.9654	–0.0220
	(–0.718)	(–0.060)	(1.012)	(–0.029)
R-squared	0.1600	0.3773	0.3185	0.1977
Adjusted R-squared	0.1446	0.3662	0.3063	0.1834
No. of observations	2,109	2,173	2,173	2,159

Note. All models include dummy variables for 20 of 21 possible math classes (*t*-statistics in parentheses).

class size and its square are significant, with increases in class size predicted to increase the percentage of the text covered, up to $N = 32.3$. Beyond that point, further increases in class size lead to predicted declines in the percentage of the text covered. However, the predicted effects of changing class size from 30 to 40 are extremely small—the percentage of the text predicted to be completed falls by only 0.25%. There is no significant non-linearity in the model of time spent on instructional activities.

We now probe the relationship between class

TABLE 2A
Ordinary Least Squares Regressions With Quadratic Terms in Class Size

	Text covered (%)	Group instruction (min/week)	Individual instruction/student (min/week)	Time on instructional activities (%)
Class size	0.2971	1.3385	−2.4141	−0.3195
	(2.17)	(3.16)	(−24.89)	(−3.892)
Class size squared	−0.0046	−0.0175	0.0342	0.0025
	(−1.78)	(−2.17)	(18.47)	(1.580)
R-squared	0.1613	0.3787	0.4124	0.1987
Adjusted R-squared	0.1455	0.3673	0.4017	0.1839
No. of observations	2,109	2,173	2,173	2,159

Note. All models include dummy variables for 20 of 21 possible math classes. These models are the same as in Table 2 with the addition of a variable for class size squared (*t*-statistics in parentheses). Sample sizes are identical to those shown in Table 2.

size and the percentage of time devoted to instructional activities more closely by modeling the subcomponents of instructional and non-instructional time. Specifically, we ask: How do teachers reallocate their time between the five activities listed earlier? We had hypothesized that a change in class size is likely to alter the marginal products of these different uses of teachers' time differentially. If the goal of teachers is to maximize student learning, they should respond to class size changes by reallocating time toward the uses of time which have experienced the largest increase in marginal productivity. Table 3 addresses this question by disaggregating instructional time (already modeled in Table 2) and non-instructional time into their constituent parts. Class size increases appear to affect three types of time use, significantly increasing the percentage of time allocated to discipline and routine administration, while significantly reducing the percentage of time allocated to review. Perhaps the most fascinating result in the table is that variations in class size do not significantly alter the percentage of class time devoted to new material. Teachers appear to guard this time for new material carefully. After an increase in class size, teachers instead respond by cutting back on review time. This finding squares with our earlier finding that class size seems to have no impact on the percentage of the textbook that teachers finish.

In a similar vein, note that total class time per week, T, has a positive effect on the percentage of class time spent on review, but a weakly negative effect on the percentage of time spent on new materials. Together, these findings suggest that math teachers, when given additional resources, either in terms of additional class time, or a fixed amount of class time but with a smaller class, respond in a very specific way. They strive to improve student understanding through additional review, rather than by covering additional new topics.

The work of Brown and Saks (1987) provides some hints as to why teachers may behave in this fashion. These authors report on a small-scale but innovative study in which observers recorded the amount of time teachers devoted to individual instruction of individual students. The results suggested that teachers spend considerably more time helping the students who are furthest behind, apparently in an attempt to narrow the achievement gap they observe between students in their classrooms. It is quite possible that math teachers in our sample respond to smaller classes by increasing review time, with a view to narrowing the gap in achievement between the top and bottom students in the classroom.

In spite of this interesting substitution toward review time, perhaps a more remarkable finding from Table 3 is that variations in class size result in very small changes in teachers' allocation of time. Consider a doubling of class size, from 20 to 40. The two largest predicted reallocations are a 3.0% decrease in time devoted to review, and a 2.5% increase in the time devoted to disciplining students. In a class with 300 minutes per week, this translates to 9 minutes less on review and 7.5 minutes more on discipline per week. Overall, teachers appear to respond to variations in class size by changing their use of time in rather minor ways. Of course, readers may want to consider the size of these predicted changes in percentage time allocations not in an absolute sense but in a relative sense. A 2.5% increase in time devoted to disci-

TABLE 3
Ordinary Least Squares Regressions for Class Time Activities

	Percent of class time spent on:				
	Discipline	Review	New material	Routine	Testing
Constant	9.2408	25.439	41.921	6.9378	16.727
	(7.608)	(10.285)	(16.216)	(7.964)	(12.649)
Class size	0.1249	–0.1478	–0.0419	0.0656	–0.0072
	(5.633)	(–3.276)	(–0.888)	(4.131)	(–0.300)
Total minutes/week	0.0008	0.0130	–0.0090	0.0000	–0.0049
	(0.523)	(4.000)	(–2.645)	(0.019)	(–2.851)
Class ability level	–1.1466	–0.4468	1.7636	–0.3481	0.2310
	(–7.271)	(–1.393)	(5.263)	(–3.082)	(1.348)
Lunch assistance at school (%)	–0.0274	0.0920	0.0004	–0.0318	–0.0201
	(–1.749)	(2.887)	(0.013)	(–2.838)	(–1.179)
Black students at school (%)	0.0389	–0.0238	–0.0517	0.0157	0.0173
	(4.265)	(–1.287)	(–2.669)	(2.406)	(1.744)
Hispanic students at school (%)	0.1291	–0.1099	–0.1307	0.0774	0.0152
	(4.570)	(–1.914)	(–2.177)	(3.824)	(0.493)
School size	–0.0010	0.0007	0.0001	–0.0006	0.0007
	(–3.189)	(1.014)	(0.170)	(–2.648)	(1.891)
Suburban	0.1706	1.4339	–2.3468	–0.0026	0.5769
	(0.446)	(1.842)	(–2.888)	(–0.010)	(1.385)
Rural	–0.3837	3.5406	–5.5327	0.6951	1.1804
	(–0.881)	(4.012)	(–6.002)	(2.238)	(2.488)
Female students in class (%)	–0.0339	0.0229	0.0475	0.0048	–0.0339
	(–3.078)	(1.020)	(2.027)	(0.609)	(–2.829)
Teacher has MA degree	–0.7895	2.2912	–0.1541	–0.9331	–0.5922
	(–2.574)	(3.668)	(–0.236)	(–4.245)	(–1.774)
Teacher experience (years)	0.1156	–0.0847	–0.1352	0.1107	–0.0123
	(1.994)	(–0.718)	(–1.097)	(2.666)	(–0.196)
Teacher experience squared	–0.0039	0.0095	–0.0036	–0.0024	0.0004
	(–2.240)	(2.664)	(–0.962)	(–1.948)	(0.231)
7th grade	–5.9903	–3.4264	9.2894	–1.2329	0.3246
	(–4.340)	(–1.218)	(3.160)	(–1.244)	(0.216)
8th grade	–3.6658	1.0490	4.2136	–0.8402	–0.8730
	(–4.658)	(0.654)	(2.514)	(–1.488)	(–1.018)
9th grade	–0.1473	–0.1825	0.4456	0.4462	–0.6323
	(–0.306)	(–0.186)	(0.435)	(1.292)	(–1.207)
11th grade	0.2876	–0.8900	0.1469	0.1679	0.0893
	(0.768)	(–1.169)	(0.185)	(0.626)	(0.219)
12th grade	–0.8170	–0.7893	–0.5742	0.7988	1.3108
	(–1.461)	(–0.694)	(–0.483)	(1.993)	(2.153)
R-squared	0.2209	0.0962	0.1892	0.0614	0.0318
Adjusted R-squared	0.2069	0.0801	0.1747	0.0446	0.0209
No. of observations	2,158	2,164	2,166	2,166	2,161

Note. All models include dummy variables for 20 of 21 possible math classes (*t*-statistics in parentheses).

pline is small in an absolute sense, but compared to the sample average of 6.1% of teachers' time spent on discipline, this change is perhaps significant.

These predicted changes in how teachers allocate their time seem smaller when compared to the natural variations in teachers' time allotment as their students become older. A teacher of a Grade 7 math class, on average, spends fully 9.3% more time on new material than does a teacher of a Grade 10 math class, while spending 6.0% less time on discipline. The message seems clear: The onset of

TABLE 3A
Ordinary Least Squares Regressions for Class Time Activities

	Percent of class time spent on:				
	Discipline	Review	New material	Routine	Testing
Class size	0.1631	–0.2437	–0.1387	0.1549	0.0640
	(2.66)	(–1.95)	(–1.06)	(3.53)	(0.96)
Class size squared	–0.0008	0.0020	0.0020	–0.0018	–0.0015
	(–0.67)	(0.82)	(0.80)	(–2.18)	(–1.15)
R-squared	0.2211	0.0965	0.1895	0.0635	0.0387
Adjusted R-squared	0.2067	0.0799	0.1746	0.0463	0.0210
No. of observations	2,158	2,164	2,166	2,166	2,161

Note. All models include dummy variables for 20 of 21 possible math classes. These models are the same as in Table 3 with the addition of a variable for class size squared (*t*-statistics in parentheses). Sample sizes are identical to those shown in Table 3.

puberty among students seems to affect teachers' time use far more than do even large variations in class size![10]

To check whether we were missing larger reactions by teachers to relatively large class sizes, we re-ran the models with a quadratic in class size. As shown in Table 3A, we found no strong evidence of increased responses to class size as class size increases. The one exception was the positive response of time spent on routine activities; here the

TABLE 4
Regressions With Fixed Effects for Teachers

	Text covered (%)	Group instruction (min/week)	Individual instruction/student (min/week)	Time on instructional activities (%)
Class size	0.0625	0.4001	–0.4937	–0.0672
	(1.28)	(2.75)	(–17.08)	(–2.39)
Total minutes/week	–0.0163	0.3298	0.0340	–0.0036
	(–3.59)	(24.26)	(12.57)	(–1.35)
Class ability level	2.3076	2.1573	0.0811	1.9128
	(6.77)	(2.11)	(0.40)	(9.67)
Female students in class (%)	0.0126	0.0291	–0.0383	1.9297
	(0.58)	(0.44)	(–2.94)	(1.52)
Teacher experience (years)	–0.0749	-2.0075	–0.4667	–0.1987
	(–0.14)	(–1.22)	(–1.43)	(–0.62)
Teacher experience squared	–0.0124	–0.0775	0.0131	–0.0067
	(–0.85)	(–1.74)	(1.48)	(–0.78)
7th grade	5.1625	-15.599	4.8254	2.9470
	(1.42)	(–1.42)	(2.21)	(1.37)
8th grade	4.9431	-9.9147	3.6869	0.8080
	(1.79)	(–1.19)	(2.23)	(0.50)
9th grade	1.8151	3.7021	–0.5419	1.0636
	(1.85)	(1.26)	(–0.93)	(1.87)
11th grade	0.0708	3.3470	0.5983	–0.3934
	(0.10)	(1.51)	(1.36)	(–0.91)
12th grade	0.0694	0.1914	0.7448	0.4974
	(0.07)	(0.06)	(1.19)	(0.81)
R-squared	0.6889	0.7778	0.8414	0.7138
No. of observations	2,109	2,173	2,173	2,159

Note. All models include dummy variables for 20 of 21 possible math classes. Sample sizes correspond to those shown in Table 2 (*t*-statistics in parentheses).

TABLE 5
Regressions for Class Time Activities, With Fixed Effects for Teachers

	Percent of class time spent on:				
	Discipline	Review	New material	Routine	Testing
Class size	0.0488	–0.0258	–0.0042	0.0107	–0.0372
	(2.27)	(–0.61)	(–0.09)	(0.73)	(–1.53)
Total minutes/week	–0.0013	–0.0028	0.0056	0.0050	–0.0063
	(–0.66)	(–0.72)	(1.34)	(3.64)	(–2.78)
Class ability level	–1.4248	0.0493	1.6459	–0.4064	0.2169
	(–9.39)	(0.17)	(5.23)	(–3.91)	(1.27)
Female students in class (%)	–0.0146	–0.0034	0.0591	–0.0085	–0.0357
	(–1.50)	(–0.18)	(2.92)	(–1.28)	(–3.25)
Teacher experience (years)	0.3080	0.3564	–0.1480	0.0972	–0.4063
	(1.26)	(0.75)	(–0.29)	(0.58)	(–1.48)
Teacher experience squared	0.0085	–0.0162	–0.0022	–0.0072	0.0119
	(1.29)	(–1.26)	(–0.16)	(–1.59)	(1.60)
7th grade	–2.9958	–4.5633	6.5575	0.1457	0.9646
	(–1.82)	(–1.41)	(1.91)	(0.13)	(0.52)
8th grade	–1.3389	–0.8568	0.8694	0.9537	0.7982
	(–1.08)	(–0.35)	(0.34)	(1.12)	(0.57)
9th grade	–1.4171	0.6104	0.1093	0.3459	0.3386
	(–3.26)	(0.72)	(0.12)	(1.16)	(0.69)
11th grade	–0.2449	1.0426	–1.3795	0.2641	–0.0714
	(–0.74)	(1.62)	(–2.01)	(1.17)	(–0.19)
12th grade	–1.5501	0.6689	–1.6133	0.6833	1.4684
	(–3.31)	(0.73)	(–1.66)	(2.13)	(2.77)
R-squared	0.7077	0.6852	0.7088	0.6775	0.6106
No. of observations	2,158	2,164	2,166	2,166	2,161

Note. All models include dummy variables for 20 of 21 possible math classes. Sample sizes correspond to those shown in Table 3 (*t*-statistics in parentheses).

effect does not strengthen but instead tails off as class size rises. But for all class sizes below 43, an increase in class size is predicted to increase the percentage of time spent on routine activities.

One possibility for which we have not fully accounted is that variations in unobserved traits of the teachers account for much of the variation in time allotments. If these unobserved variations in teacher traits were correlated with the class sizes the teachers received, our estimated effects of class size could be biased in either direction. Fortunately, many of the teachers teach more than one class in the sample. This panel nature of the data allows us to add a dummy variable for each teacher. This fixed effect model removes all variation between teachers in the dependent and explanatory variables. The effect of variations in class size will be identified solely by variations for the individual teacher in the classes he or she teaches. Of course, when estimating such a specification, many teachers effectively drop out of the new sample since they are represented only once in the original sample. Other teachers provide some identifying information, but very little, since they teach only two classes. Still, this provides a useful check that the coefficients we obtained earlier, all of which point to the notion that class size has little effect on teacher behavior, are robust to the presence of unobserved fixed teacher effects.

Tables 4 and 5 repeat the models from Tables 2 and 3 with teacher fixed effects. The coefficients on class size drop somewhat, but these models continue to indicate that reductions in class size lead to a significant substitution away from group instruction towards individual instruction. Table 5 shows that the coefficients on class size in the models of the other set of time allocation variables drop considerably. However, the signs of the predicted effects are identical, and in the case of the percentage of class time spent on student discipline, the result continues to hold that one side effect of larger classes is more time devoted to student discipline. These models do not necessarily improve on the earlier OLS models, because they ignore all variation in class size across teachers. This reduces the precision of the estimates, but potentially removes bias due to unobserved traits of the teacher. The results seem to indicate that, as before, variations in class size do not have large effects on teacher behavior.

Checking for Interactions Between Class Size and Traits of the Teacher, Student Body, or Class

Given the strength of the general result that variations in class size don't seem to alter math teachers' time allocation substantially, it becomes crucial to check whether this is true for all teachers, the student bodies at all schools, and math classes at varying achievement levels. We probe these questions by modifying the regressions in Tables 2 and 3 by adding an interaction between class size and other regressors, one at a time.

First, we examine whether more highly educated teachers respond differently to variations in class size. Table 6 shows the coefficients on class size, a dummy variable for whether the teacher holds a Master's degree, and an interaction between the two. In this and subsequent tables, the top panel shows the modified versions of the regressions in Table 2, and the bottom panel shows the modified versions of the models in Table 3. The table reveals that teachers at both levels (Master's or less) for the most part show no significant difference in how they alter their time allocation when class size changes. There are two exceptions: Teachers at the Master's level are mainly responsible for the result that a rise in class size increases time spent on group instruction. On the other hand, our earlier result that a rise in class size leads to a greater percentage of time devoted to routine seems to derive solely from the less highly educated teachers. In our judgement, for neither level of teacher education do teachers seem to respond to class size variations with large reallocations of their time.

Table 7 extends the models in Tables 2 and 3 by adding an interaction between class size and teachers' years of experience. Several of the interaction terms are significant at 5% or 10% in this table. However, as teacher experience grows, most of the variations in teacher response to class size changes

TABLE 6
Regressions With Interactions Between Class Size and a Dummy for Whether Teacher Holds a Master's Degree

	Text covered (%)	Group instruction (min/week)	Individual instruction/student (min/week)	Time on instructional activities (%)
Class size	0.0210	0.0614	–0.6846	–0.2162
	(0.319)	(0.303)	(–13.710)	(–5.496)
Teacher MA degree	–1.0350	–18.113	3.0502	0.6265
	(–0.467)	(–2.673)	(1.826)	(0.477)
Teacher MA × class size	0.0987	0.8567	–0.1214	0.0360
	(1.118)	(3.161)	(–1.818)	(0.684)
R-squared	0.1605	0.3802	0.3195	0.1979
Adjusted R-squared	0.1447	0.3689	0.3071	0.1831
No. of observations	2,109	2,173	2,173	2,159

	Percent of class time spent on:				
	Discipline	Review	New material	Routine	Testing
Class size	0.1089	–0.1656	–0.0391	0.1007	–0.0096
	(3.706)	(–2.768)	(–0.625)	(4.790)	(–0.300)
Teacher MA degree	–1.5610	1.4312	–0.0191	0.7637	–0.7067
	(–1.591)	(0.717)	(–0.009)	(1.087)	(–0.662)
Teacher MA × class size	0.0325	0.0362	–0.0057	–0.0714	0.0048
	(0.828)	(0.453)	(–0.068)	(–2.543)	(0.113)
R-squared	0.2212	0.0963	0.1892	0.0643	0.0381
Adjusted R-squared	0.2068	0.0797	0.1744	0.0471	0.0204
No. of observations	2,158	2,164	2,166	2,166	2,161

Note. t-statistics in parentheses.

TABLE 7
Regressions With Interactions Between Class Size and Teacher Experience

	Text covered (%)	Group instruction (min/week)	Individual instruction/student (min/week)	Time on instructional activities (%)
Class size	0.0157	0.0567	−1.0540	−0.1720
	(0.170)	(0.198)	(−15.063)	(−3.100)
Teacher experience (years)	−0.2662	−1.9549	−0.4590	−0.1854
	(−1.458)	(−3.482)	(−3.341)	(−1.703)
Teacher experience squared	0.0040	0.0328	−0.0026	0.0063
	(1.022)	(2.708)	(−0.882)	(2.699)
Teacher experience × class size	0.0036	0.0285	0.0208	−0.0018
	(0.689)	(1.758)	(5.246)	(−0.566)
R-squared	0.1602	0.3782	0.3271	0.1979
Adjusted R-squared	0.1444	0.3669	0.3148	0.1831
No. of observations	2,109	2,173	2,173	2,159

	Percent of class time spent on:				
	Discipline	Review	New material	Routine	Testing
Class size	0.0468	0.0549	−0.1365	0.0970	−0.0933
	(1.132)	(0.652)	(−1.549)	(3.265)	(−2.071)
Teacher experience (years)	−0.0115	0.2447	−0.2891	0.1616	−0.1524
	(−0.142)	(1.483)	(−1.673)	(2.776)	(−1.726)
Teacher experience squared	−0.0038	0.0093	−0.0035	−0.0025	0.0005
	(−2.194)	(2.616)	(−0.938)	(−1.972)	(0.274)
Teacher experience × class size	0.0052	−0.0136	0.0064	−0.0021	0.0058
	(2.233)	(−2.848)	(1.271)	(−1.248)	(2.261)
R-squared	0.2227	0.0997	0.1898	0.0621	0.0404
Adjusted R-squared	0.2084	0.0831	0.1750	0.0449	0.0228
No. of observations	2,158	2,164	2,166	2,166	2,161

Note. t-statistics in parentheses.

are quite small. We find one important exception to this rule. Table 3 illustrates that teachers react to larger classes by cutting back on review time. Table 7 shows that this reaction to hikes in class size grows significantly among more highly experienced teachers. But even here, the overall effect of class size changes is modest. For instance, a math teacher with 25 years of experience (about 1 standard deviation above the mean) is predicted to respond to an increase of class size of 10 by reducing the percentage of time devoted to review by 10 [25(-0.0136) + 0.0549] = -2.85%. This is a meaningful drop, but it is a modest reduction from the sample mean of 30.66% of class time.

Tables 8 and 9 add interactions between traits of the overall student body at the school and traits of the students in the specific class, respectively. Table 8 interacts class size with the percentage of students at the school receiving full lunch assistance. For the most part, the interaction terms are insignificant. At schools with a larger percentage of students receiving assistance, math teachers are slightly less likely to react to class size increases by reducing the minutes of individual instruction per week. Again, the change is rather small: Consider two schools, one with no children on full lunch assistance, and a second with 30% on full lunch assistance. These schools are approximately 1 standard deviation below and above the mean, respectively. At the first school, an increase in class size by one reduces individual instruction by 0.826 minutes per week. At the second school, the same change in class size reduces time spent on individual instruction by 0.682 minutes per week.[11]

Turning to the students in the actual classroom, we next interacted class size in the models in Tables 2 and 3 with the teacher's estimate of class ability, on a 1–5 scale. Table 9 reveals some highly significant and meaningful variations across class ability levels. Recall that the teacher's estimate of

TABLE 8
Regressions With Interactions Between Class Size and Percentage of Students in School Receiving Lunch Assistance

	Text covered (%)	Group instruction (min/week)	Individual instruction/student (min/week)	Time on instructional activities (%)
Class size	0.0943	0.6321	–0.8261	–0.2203
	(1.453)	(3.160)	(–16.800)	(–5.697)
Lunch assistance (%)	–0.0507	0.4535	–0.1625	0.0410
	(–0.758)	(2.206)	(–3.217)	(1.031)
Lunch assistance (%) × class size	–0.0015	–0.0088	0.0048	0.0013
	(–0.605)	(–1.175)	(2.595)	(0.877)
R-squared	0.1602	0.3777	0.3206	0.1980
Adjusted R-squared	0.1443	0.3664	0.3082	0.1833
No. of observations	2,109	2,173	2,173	2,159

	Percent of class time spent on:				
	Discipline	Review	New material	Routine	Testing
Class size	0.1357	–0.1220	–0.1005	0.0768	0.0034
	(4.695)	(–2.076)	(–1.637)	(3.711)	(0.108)
% Lunch assistance	–0.0128	0.1271	–0.0794	–0.0166	–0.0057
	(–0.431)	(2.104)	(–1.259)	(–0.782)	(–0.176)
% Lunch assistance × class size	–0.0006	–0.0015	0.0034	–0.0007	–0.0006
	(–0.581)	(–0.684)	(1.490)	(–0.842)	(–0.527)
R-squared	0.2210	0.0964	0.1901	0.0617	0.0383
Adjusted R-squared	0.2067	0.0798	0.1752	0.0445	0.0206
No. of observations	2,158	2,164	2,166	2,166	2,161

Note. t-statistics in parentheses.

class ability varies between 1 (much lower than average) and 5 (much higher than average). There are three cases in which the sign of the derivative of a time allocation variable with respect to class size differs between the bottom and top ability groups; in a fourth case the sign never changes but varies significantly with class ability, both in a statistical and a policy sense. Consider first the pair of variables measuring group instruction time and individual instruction time per student. Table 2 showed that on average an increase in class size leads to more time spent on group instruction and less time on individual instruction. Table 9 suggests that this pattern is quite strong in the classes with the lowest achievement, and zero or even reversed in the top classes. Similarly, the negative overall impact of class size on the percentage of time spent on instructional activities appears to weaken significantly in the top classes. The bottom panel of the table shows that the variation in how teachers change instructional time can be accounted for by sharp variations in the percentage of time spent on review. The figures suggest that for the top two levels of class ability, teachers do not adjust time spent on review when class size changes much, if at all.

The size of these interactions is far larger than any found in the earlier tables. At the bottom class ability levels, a rise in class size is predicted to change teachers' time allocations by relatively large amounts. It becomes important to know what is driving these results, especially given that our coding of teachers' responses to the question about class ability into a 1–5 scale is rather ad hoc. What is behind the large interaction terms in some of the columns of Table 9? Is it variations in teacher responses in the bottom classes, the top classes, or both? Accordingly, Tables 10 and 11 report the class size coefficient in replications of the models in Tables 2 and 3 that were run on each of the class ability subsamples. (We had to drop certain dummy variables for type of math class in the regressions on these subsamples to prevent perfect collinearity. Apart from that, the specifications conform exactly to those in Tables 2 and 3.)

Each row in Table 10 lists the coefficient and corresponding *t*-statistic from a regression with the dependent variable stated at the top of the column, and with the class ability subsample indicated in the first column. We lose some precision in these subsamples, and at the same time risk selectivity bias. Nevertheless, the results support those in Table 9. The impact of changes in class size on minutes of group instruction and on individual instruction per student are far larger for classes judged to be of lower ability (1 and 2). In fact, for the top three ability groups there is no significant evidence that teachers alter total group minutes at all. In contrast, for the bottom two groups an increase in class size by 10 increases predicted minutes of group instruction per week by 27 and 14 minutes for the first and second groups, respectively. These shifts in time allocation might have substantial impacts on these students. Similarly, and perhaps more importantly, the negative impact of increases in class size on the proportion of class time devoted to instructional activities occurs most strongly in classes taught to less advanced students. The coefficient on class size is not always significant at the 5% level in these regressions, but the coefficient rises (towards zero) monotonically as the level of ability of the class increases. An increase in class size of 10 is predicted to reduce the percentage of time devoted to instructional activities by 3% in the bottom classes, compared to just 0.6% in the top classes. Neither of these changes is very large, but the pattern is noteworthy.

Table 11 shows how the class size coefficient varies when the five components of total class time are modeled by subsample. Here monotic relationships to class ability level do not emerge, but overall the results suggest that lower ability classes see a higher reduction in review time when class size increases.

As a final check on whether stronger effects of class size on teachers' time allocation might emerge in certain types of classes, we examine whether middle school and high school teachers respond in the same way. Since so much else can vary between the two types of schools, we ran separate

TABLE 9
Regressions With Interactions Between Class Size and Class Ability

	Text covered (%)	Group instruction (min/week)	Individual instruction/student (min/week)	Time on instructional activities (%)
Class size	0.2142	1.9759	−1.8422	−0.3425
	(1.639)	(4.954)	(−19.367)	(−4.425)
Class ability level	3.2227	10.324	−7.7402	0.5317
	(3.490)	(3.654)	(−11.487)	(0.972)
Class ability × class size	−0.0441	−0.4573	0.3360	0.0440
	(−1.200)	(−4.057)	(12.497)	(2.013)
R-squared	0.1606	0.3821	0.3650	0.1993
Adjusted R-squared	0.1448	0.3708	0.3534	0.1845
No. of observations	2,109	2,173	2,173	2,159

	Percent of class time spent on:				
	Discipline	Review	New material	Routine	Testing
Class size	0.1883	−0.5424	0.1078	0.1174	0.0961
	(3.260)	(−4.622)	(0.877)	(2.832)	(1.528)
Class ability level	−0.6974	−3.2342	2.8210	0.0174	0.9612
	(−1.704)	(−3.898)	(3.243)	(0.059)	(2.161)
Class ability × class size	−0.0194	0.1206	−0.0458	−0.0158	−0.0316
	(−1.189)	(3.640)	(−1.317)	(−1.351)	(−1.779)
R-squared	0.2214	0.1018	0.1899	0.0622	0.0396
Adjusted R-squared	0.2071	0.0853	0.1750	0.0450	0.0219
No. of observations	2,158	2,164	2,166	2,166	2,161

Note. t-statistics in parentheses.

TABLE 10
Class Size Coefficients From Regressions in Table 2 Repeated on Subsamples by Teacher's Estimate of Class Ability

Regression subsample	Text covered (%)	Group instruction (min/week)	Individual instruction/student (min/week)	Time on instructional activities (%)
Class Ability = 1	0.1057	2.7438	−1.9158	−0.3093
	(0.353)	(4.045)	(−6.143)	(−1.677)
Class Ability = 2	0.4159	1.4096	−1.3247	−0.2753
	(3.080)	(3.810)	(−8.510)	(−3.252)
Class Ability = 3	0.0631	0.1650	−0.6234	−0.2325
	(0.673)	(0.562)	(−11.993)	(−3.849)
Class Ability = 4	−0.0649	−0.2523	−0.2908	−0.1862
	(−0.774)	(−0.832)	(−15.407)	(−3.801)
Class Ability = 5	−0.0145	−0.1237	−0.3867	−0.0647
	(−0.138)	(−0.359)	(−16.491)	(−1.276)

Note. Each row refers to regression coefficient and *t*-statistic (in parentheses) on the class size variable from a set of models with the dependent variable listed at the top of the column and performed using the class ability subsample indicated in column 1.

regressions for students in Grades 7 and 8 and for Grades 9 through 12. The results appear in the Appendix in Tables A1 and A2, which reproduce the models in Table 2. Patterns observed earlier in the pooled data set appear to hold in both types of schools, although middle school math teachers seem to respond more strongly to variations in class size than do their counterparts in high schools. Most strikingly, an increase in class size by 10 students in middle schools is predicted to increase group instruction minutes by 17.6 minutes per week, compared to an increase of only 2.9 minutes in high schools. Other differences are quite muted compared to this gap, but tend to tell the same story: Middle school teachers react more strongly to changes in class size than do high school teachers.

5. Conclusion

In this paper we have sought to open the black box of CSR by studying how variations in class size affect the amount of material covered by the teacher, and more specifically and fundamentally, how teachers reallocate their time when class size changes. The largest effects we found were that a reduction in class size induces teachers to devote less time to group instruction and more time to

TABLE 11
Class Size Coefficients From Regressions in Table 3 Repeated on Subsamples by Class Ability

	Percent of class time spent on:				
Regression subsample	Discipline	Review	New material	Routine	Testing
Class Ability = 1	0.1909	−0.2177	−0.0345	0.0756	−0.0571
	(1.170)	(−1.123)	(−0.168)	(0.913)	(−0.544)
Class Ability = 2	0.1616	−0.4702	0.0715	0.1150	0.1278
	(2.556)	(−4.488)	(0.692)	(2.526)	(2.439)
Class Ability = 3	0.1602	−0.1283	−0.0290	0.0670	−0.0747
	(3.415)	(−1.427)	(−0.314)	(2.262)	(−1.541)
Class Ability = 4	0.1291	−0.1254	−0.0482	0.0490	−0.0144
	(3.826)	(−1.487)	(−0.524)	(1.699)	(−0.302)
Class Ability = 5	0.0452	0.2228	−0.1990	0.0401	−0.0854
	(1.737)	(2.117)	(−1.728)	(1.159)	(−1.567)

Note. Each row refers to regression coefficient and *t*-statistic (in parentheses) on the class size variable from a set of models with the dependent variable listed at the top of the column and performed using the class ability subsample indicated in column 1.

individual instruction. However, in earlier work (Shkolnik & Betts, 1998) we use a similar data extract and conclude that, if anything, teachers could make CSR considerably more effective if they did not reduce group instruction to the extent that they do.

This paper reveals that class size variations induce little change in how teachers allocate their time between new material, review, discipline, routine tasks, and testing. The fact that teachers reallocate their time to such a small extent may explain why it has been so hard in most past research to identify a positive and significant impact of CSR on student achievement.[12]

We uncovered a fascinating pattern in the data. When class size is reduced, teachers do not increase the proportion of time spent on new material. Instead, they allocate more time to review activities. We found similar behavior when total class time per week was increased: Teachers reacted not by adding on more material, but by expanding review time instead. Similarly, neither reductions in class size nor increases in minutes of class per week significantly affected the percentage of the textbook covered by teachers. Put another way, teachers appear to react to *increases* in class size by guarding the time they devote to new material, while cutting back on review time.

But even here, the size of the time reallocations was extremely small. A halving of class size from 40 to 20 is predicted to lead to a rise of only 3 percentage points in the percentage of time devoted to review.

We undertook extensive specification tests to examine whether variations in class size mattered more in some types of classes or schools, or with certain types of teachers more than others. We uncovered a number of systematic variations. By far the most important of these is that the tendency of teachers to substitute group instruction for individual instruction as class size increases is far stronger in classes identified by the teacher as "below average" or "much below average" than in more able classes. We also found evidence that middle school teachers react more strongly to changes in class size than do high school teachers, but the differences are not as impressive. On average, and for most types of teachers, student bodies, and class abilities, teachers' reallocation of class time as class size changes appears to be quite modest.

Overall, this set of findings raises the possibility that CSR might become more effective if teachers adjusted their teaching styles more radically so as to take advantage of smaller classes. Further research is required to identify exactly what those changes in teaching style should entail.

APPENDIX

The group instruction and individual instruction per student variables were created using teachers' responses when they were asked how much time they allocated to the following five activities: lecturing, leading discussion, working in small groups, doing seatwork, and providing individual instruction.[13] Possible responses include: (1) *None*, (2) *30 minutes*, (3) *1 hour*, (4) *2 hours*, and (5) *more than 2 hours*. We modify these variables to time in minutes: (1) *0 minutes*, (2) *30 minutes*, (3) *60 minutes*, (4) *120 minutes*, and (5) *150 minutes*. To construct the group instruction variable, we combine the variables for time spent on lecture and discussion, both of which have the teacher in front of the class leading the entire group. Individual per student instruction is created by combining minutes spent on individual instruction, seatwork, and in small groups (where students receive individual instruction or tutoring) and dividing this sum by class size.

The variables on the percentage of the textbook covered are obtained from the question, "What percentage of the textbook will you cover this year?" (The questionnaire is sent out in spring of the school year.)

We derive the variables indicating the percentage of time devoted to various topics from the following set of questions:

> Overall, what percentage of your classroom time is spent in each of the following:
>
> (1) Daily routines (such as set up, clean up, passing out materials, taking attendance, announcements, breaks).
> (2) Getting students to behave.
> (3) Presenting new material.
> (4) Review or student practice of skills.
> (5) Testing or other forms of evaluation.

We refer to these variables as the percentage of class time spent on routine, discipline, new material, review and testing, respectively.

We added dummy variables to control for 20 of 21 types of math courses taught in the LSAY schools. The coding of this information varies by grade level, but can be manipulated to provide an entirely consistent set of course types. These are:

Low 7th grade
Average 7th grade
High 7th grade
Low 8th grade
Average 8th grade

- High 8th grade
- Basic (9–12)
- Vocational
- Consumer
- Geometry
- Geometry Honors
- Pre-Algebra
- Algebra I
- Algebra I Honors
- Algebra II
- Algebra II Honors
- Trigonometry
- Trigonometry Honors
- Analytical Geometry/Pre-Calculus
- Calculus
- Statistics/Probability

TABLE A1
Seventh and Eighth Grade Ordinary Least Squares Regressions

	Text covered (%)	Group instruction (min/week)	Individual instruction/student (min/week)	Time on instructional activities (%)
Constant	51.044	17.565	16.004	88.510
	(7.837)	(0.923)	(4.193)	(19.703)
Class size	0.2579	1.7592	–0.9613	–0.2888
	(2.208)	(5.324)	(–14.509)	(–3.681)
Total minutes/week	0.0070	0.3196	0.0395	0.0063
	(0.950)	(14.785)	(9.122)	(1.234)
Class ability level	3.7019	1.7473	–0.7361	2.3253
	(4.305)	(0.700)	(–1.471)	(3.936)
Lunch assistance at school (%)	–0.1405	–0.3915	0.0132	–0.0351
	(–2.268)	(–2.250)	(0.379)	(–0.857)
Black students at school (%)	0.0862	0.1724	–0.0058	0.0374
	(1.967)	(1.410)	(–0.235)	(1.295)
Hispanic students at school (%)	0.0569	0.0272	0.0442	–0.1093
	(0.632)	(0.106)	(0.859)	(–1.796)
School size	0.0062	–0.0278	0.0044	–0.0008
	(2.084)	(–3.283)	(2.588)	(–0.414)
Suburban	0.2809	–8.3338	2.0570	–1.7462
	(0.143)	(–1.469)	(1.808)	(–1.304)
Rural	0.1145	4.3381	1.0364	–0.2262
	(0.057)	(0.737)	(0.878)	(–0.163)
Female students in class (%)	0.0856	0.1649	0.0707	0.0059
	(1.484)	(0.980)	(2.097)	(0.150)
Teacher has MA degree	1.4483	5.5264	–0.3147	0.9898
	(0.922)	(1.223)	(–0.347)	(0.929)
Teacher experience (years)	–0.4043	–2.5987	–0.0411	0.3600
	(–1.217)	(–2.785)	(–0.220)	(1.635)
Teacher experience squared	0.0169	0.0702	0.0006	–0.0154
	(1.548)	(2.254)	(0.097)	(–2.101)
7th grade	–2.3952	1.3215	2.3259	0.8829
	(–0.801)	(0.150)	(1.314)	(0.425)
R-squared	0.2463	0.4507	0.5241	0.2682
Adjusted R-squared	0.2002	0.4197	0.4972	0.2265
No. of observations	382	412	412	409

Note. All models include dummy variables for 20 of 21 possible math classes (*t*-statistics in parentheses).

TABLE A2
Ninth Through Twelfth Grade Ordinary Least Squares Regressions

	Text covered (%)	Group instruction (min/week)	Individual instruction/student (min/week)	Time on instructional activities (%)
Constant	80.394	39.655	14.299	84.152
	(26.333)	(4.160)	(5.871)	(47.748)
Class size	0.0172	0.2971	−0.7047	−0.1592
	(0.309)	(1.713)	(−15.901)	(−4.961)
Total minutes/week	0.0012	0.2959	0.0333	−0.0021
	(0.294)	(23.291)	(10.255)	(−0.900)
Class ability level	1.7922	−0.7956	0.1674	1.3340
	(4.605)	(−0.655)	(0.539)	(5.935)
Lunch assistance at school (%)	−0.0917	0.5793	−0.0867	0.1047
	(−2.056)	(4.187)	(−2.454)	(4.053)
Black students at school (%)	−0.0545	−0.1206	0.0301	−0.0869
	(−2.268)	(−1.617)	(1.580)	(−6.246)
Hispanic students at school (%)	0.3320	−0.9933	0.2203	−0.1891
	(3.155)	(−3.013)	(2.616)	(−3.090)
School size	−0.0018	0.0124	−0.0008	0.0015
	(−2.300)	(4.991)	(−1.299)	(3.332)
Suburban	−0.4180	4.1992	0.0281	−0.2542
	(−0.431)	(1.387)	(0.036)	(−0.452)
Rural	−2.1177	−9.7945	1.9305	−0.9462
	(−1.911)	(−2.825)	(2.180)	(−1.462)
Female students in class (%)	−0.0351	−0.0592	−0.0024	0.0353
	(−1.295)	(−0.698)	(−0.111)	(2.239)
Teacher has MA degree	1.0771	1.7126	0.2496	1.5168
	(1.408)	(0.717)	(0.409)	(3.426)
Teacher experience (years)	−0.1184	−1.0244	0.0515	−0.2992
	(−0.820)	(−2.267)	(0.446)	(−3.581)
Teacher experience squared	0.0020	0.0261	−0.0033	0.0089
	(0.460)	(1.953)	(−0.961)	(3.600)
9th grade	0.1158	3.0428	0.2188	−0.4351
	(0.106)	(0.895)	(0.252)	(−0.691)
11th grade	−0.3250	0.1296	1.4326	−0.6515
	(−0.388)	(0.049)	(2.132)	(−1.338)
12th grade	−0.9976	−0.5050	1.1155	−0.0124
	(−0.792)	(−0.128)	(1.108)	(−0.017)
R-squared	0.1600	0.3436	0.2905	0.1942
Adjusted R-squared	0.1451	0.3322	0.2782	0.1801
No. of observations	1,727	1,761	1,761	1,750

Note. All models include dummy variables for 20 of 21 possible math classes (*t*-statistics in parentheses).

Notes

Betts acknowledges support from the Social Sciences and Humanities Research Council of Canada. This research was also supported in part by a Dissertation Grant to Shkolnik from the American Educational Research Association, which receives funds for its "AERA Grants Program" from the National Science Foundation and the National Center for Education Statistics (U.S. Department of Education) under NSF Grant #RED-9452861. The opinions stated here reflect those of the authors and do not necessarily reflect those of the granting agencies.

[1] For evidence of declining pupil-teacher ratios over the last few decades, see for instance Hanushek and Rivkin (1997).

[2] For a competing view of the test-score literature, see Hedges and Greenwald (1996).

[3] An isoquant curve traces out the combinations of any two or more inputs (here, the two types of instruction) which give rise to the same level of output (student achievement). These combinations must be the most efficient ones—i.e., any point on an isoquant curve shows the

minimum quantities of the inputs needed to produce the given output (Pearce, 1992).

[4] This curved shape reflects the notion that if a teacher increases the level of one input, say, group instruction, while holding the other input constant, eventually there will be diminishing returns to increasing group instruction any further. The slope of the isoquant shows the tradeoff between the two inputs. So if I is very low, and G is very high, the isoquant is very steep, indicating that it would take a very large increase in G to keep student achievement constant after reducing I by one more unit. Similarly, at low G and high I, the isoquant becomes quite flat, indicating that the few remaining minutes of group instruction are now rather precious relative to individual instruction.

[5] Formally, the teacher might increase individual instruction per student (I) while reducing group instruction (G) if the positive income effect on G is outweighed by a negative substitution effect when class size is reduced.

[6] The LSAY did not include students enrolled in special education, but large variations in teachers' perceptions of each class's ability or achievement level remain.

[7] The question reads, for example, for seventh-grade teachers: "How would you rate the average academic ability of the students in this class compared to all 7th graders in your school?" Answers include: much higher than average, somewhat higher, about average, somewhat lower, and much lower than average.

[8] We also could have modeled total individual instructional time ($I \times N$) instead of average individual instruction per student. However, since $I = T - G$, the coefficients in such a regression will equal the negative of those displayed in the model of G, while the coefficient on T would have been 1 minus the coefficient on total time T in the model of G.

[9] Again, our regressor in column 3, $(T-G)/N$, leads to a positive second derivative with respect to class size N even if total group instruction G does not depend on N.

[10] Interestingly, though, the impact of class size on time use is similar in magnitude to the impact of variations in the traits of the student body. In the thought experiment considered above, an increase of 20 students in the class approximates a 3 standard deviation increase in class size (see Table 1). A 3 standard deviation increase in the percentage of the student body that is Black (about a 60% increase) is predicted to reduce time spent on new material by 3.1%, while increasing time spent on discipline by 2.3%. A 3 standard deviation increase in the percentage of Hispanic students in the school (an 18% increase) is predicted to decrease time on new material by 2.4% while increasing time spent on discipline by 2.3%.

[11] We also repeated the models in Tables 2 and 3 with interactions between class size and the percentage of the student body that was Black or that was Hispanic. The results, in terms of statistical significance and the size of the interactions, were quite similar to those reported in Table 8 for the interaction between class size and lunch assistance.

[12] See also Betts (1997), which shows that math teachers with larger classes tend to assign slightly more homework each week. Because homework assigned is a highly significant predictor of gains in test scores, such compensating behavior on the part of teachers provides another reason why it may be hard to find a strong link between gains in test scores and class size.

[13] The question reads, "About how much classroom time do you spend on each of the following with this class during a typical week?" The categories include lecturing to the class; leading discussions; having students work in small groups or a laboratory; having students do seatwork on homework, workbook, or text assignments; and providing individualized instruction.

References

Betts, J. R. (1996). Is there a link between school inputs and earnings? Fresh scrutiny of an old literature. In G. Burtless (Ed.), *Does money matter? The effect of school resources on student achievement and adult success* (pp. 141–91). Washington, DC: Brookings Institution.

Betts, J. R. (1997). *The role of homework in improving school quality.* (Discussion Paper 96–16). University of California, San Diego, Department of Economics.

Betts, J. R., & Shkolnik, J. L. (in press). The effects of ability grouping on student achievement and resource allocation in secondary schools. *Economics of Education Review.*

Brown, B. W., & Saks, D. H. (1987). The microeconomics of the allocation of teachers' time and student learning. *Economics of Education Review, 6,* 319–332.

Finn, J. D., & Achilles, C. M. (1990). Answers and questions about class size: A statewide experiment. *American Educational Research Journal, 27,* 557–577.

Hanushek, E. A. (1996). School resources and student performance. Chapter 2. In G. Burtless (Ed.), *Does money matter? The effect of school resources on student achievement and adult success* (pp. 43–73). Washington, DC: Brookings Institution.

Hanushek, E. A., & Rivkin, S. G. (1997). Understanding the twentieth-century growth in U.S. school spending. *Journal of Human Resources, 32,* 35–68.

Hedges, L. V., & Greenwald, R. (1996). Have times changed? The relation between school resources and student performance. Chapter 3. In G. Burtless (Ed.), *Does money matter? The effect of school resources on student achievement and adult success* (pp. 74–92). Washington, DC: Brookings Institution.

Krueger, A. B. (1997). *Experimental estimates of education production functions* (Working Paper No. 6051). Cambridge, MA: National Bureau of Economic Research.

Mosteller, F. (1995). The Tennessee study of class size in the early school grades. *The Future of Children: Critical Issues for Children and Youths, 5,* 113–127.

Murnane, R. J. (1991). Interpreting the evidence on "Does Money Matter?" *Harvard Journal on Legislation, 28,* 457–464.

Pearce, D. W. (Ed.). (1992). *The MIT Dictionary of Modern Economics.* Cambridge, MA: The MIT Press.

Shkolnik, J. L., & Betts, J. R. (1998). *The effects of class size on teacher time allocation and student achievement.* Unpublished manuscript, Department of Economics, University of California, San Diego.

Authors

JULIAN R. BETTS is an associate professor at the Department of Economics, University of California-San Diego, La Jolla, CA 92093-0508, and a senior fellow at the Public Policy Institute of California, 500 Washington Street, Suite 800, San Francisco, CA 94111; e-mail: **jbetts@ucsd.edu**. He specializes in the economics of education.

JAMIE L. SHKOLNIK is a research scientist at the National Opinion Research Center, 1350 Connecticut Avenue NW, Suite 500, Washington, DC 20036; e-mail: **shkolnik-jamie@norcmail.uchicago.edu**. Her area of specialization is the economics of education.

Manuscript received January 6, 1999
Revision received February 17, 1999
Accepted February 18, 1999