Would Reductions in Class Size Raise Minnesota Students' Test Scores? Evidence from Minnesota's Elementary Schools

by Hyunkuk Cho, Paul Glewwe, and Melissa Whitler

Policymakers, parents, school principals, and pundits are all concerned about how much, or in some cases how little, students learn in America's schools. Despite substantial research in recent years, much is unknown about the impacts of specific education policies on student learning. One education policy that has received much attention is reductions in class size. Intuitively, smaller classes should allow teachers to provide more attention to each student and to reduce time spent disciplining disruptive students, and thus should increase students' learning. Indeed, there appears to be a consensus among parents, teachers, and school administrators that small classes improve students' academic achievement, especially among elementary school students.

Yet this consensus is not supported by academic research, which has found conflicting evidence on the impact of class size on learning. The basic problem is that students in small and large classes may differ in many other ways. If one has data on these differences, a variety of statistical methods can be used to account for these differences. But if one does not have data on these differences, then almost all standard estimation methods will lead to biased estimates of the impact of class size on student learning.

Perhaps the best method to measure the impact of class size reductions (and many other types of education policies) on student learning is to randomly assign some children to small classes and other children to large classes, and compare the educational outcomes of interest across the two groups. Random assignment ensures that, on average, the two groups of students have the same observed and unobserved characteristics. In the United States, only one study has implemented this type of research on a large scale: Project STAR (Student/Teacher Achievement Ratio Experiment), which was conducted in Tennessee from 1985 to 1989.

Early results from Project STAR suggested that students in small classes scored higher on mathematics and reading tests than did students in regular-size classes. These results influenced policies in other states.

For example, California's decision to reduce class sizes in grades K–3 to less than 20, which was first implemented in the 1996–1997 school year, was directly influenced by the results from Project STAR. Also in the 1996–1997 school year, Wisconsin initiated a program to reduce class size in grades K–3 to less than 15. The U.S federal government also became involved when it implemented the Class-Size Reduction Program in fiscal year 1999, which provides funds to states for hiring new teachers to reduce class size in grades 1–3.

Should Minnesota follow the lead of these other states and reduce class size in the first grades of elementary school? Even if reducing class size would increase test scores, class size reductions may not be a wise policy choice because class size reductions are very expensive. That is, there may be other policies that increase learning by an equivalent amount yet at a lower cost. California spent $11 billion and Wisconsin spent $463 million from 1996 to 2005 to attain their class size reduction goals. Before committing such large amounts of money, a clearer idea of the likely benefits in terms of child learning is needed. This article provides estimates of the likely effect of class size reductions in Minnesota, based on test score data from 1997–1998 to 2004–2005 for children in grades 3 and 5 in Minnesota public schools. The research presented here was supported by a Faculty Interactive Research Program grant from CURA.

Photo © Corbis, 2007

Intuitively, smaller classes should allow teachers to provide more attention to each student, thus increasing students' learning. However, studies have found conflicting evidence on the impact of class size on learning, as well as the degree of such impacts.

A Review of the Literature on the Impact of Class Size on Learning

Ever since the publication in 1966 of *Equality of Educational Opportunity* (also known as the Coleman Report), social science researchers have attempted to estimate the influence of a variety of factors on student learning, including class size. Yet the research up to the mid-1990s was of variable quality because of inadequate data and serious estimation problems. The uneven quality of the research led to a lack of consensus on the impact of class size on learning. In a 1997 review of the literature from the late 1960s to the early 1990s published in *Educational Evaluation and Policy Analysis,* Eric Hanushek reported that, of 277 studies that attempted to estimate the impact of class size on student performance, 15% found an unexpected statistically significant positive effect (that is, increased class size increased student performance), 13% found the expected statistically significant negative effect (increased class size decreased student performance), and the remaining 72% found no statistically significant effect whatsoever.

In the last 10 years, much more careful analyses have been done of the impact of class size on student learning. The best studies focus on removing or at least reducing likely sources of bias in estimates of that impact. Almost all sources of bias arise because class size is correlated with unobserved student, parent, or school variables that directly affect student learning. For example, parents who are very concerned about their children's education may be more likely to move to areas where schools have small classes, which would lead to a situation in which students in schools with small class sizes have, on average, parents who provide more support at home for their children's education. This will cause overestimation of the impact of class size reductions on student learning if estimation methods do not account for this relationship between class size and parental aspirations (which will be difficult to do if there are no data on parental aspirations or parental support at home, which is usually the case). The intuition here is that part of the positive association between small class size and learning is really caused by the fact that parents who make extra efforts for their children's education are also more likely to enroll their children in schools with small classes. A second problem is that parents may enroll their children in schools that they perceive to be of high quality, thus increasing class sizes in schools perceived to be of high quality. If some or all of the school quality characteristics that parents use to form their perceptions are not in the data set, this behavior will lead to a situation in which schools with large class sizes also tend to be of better quality, which will cause underestimation of the impact of reductions in class size on learning.[1] A third problem is that educators may assign students to classes of different sizes depending on their abilities; estimates will be biased if student ability is not taken into account, and the direction of bias will depend on whether high-ability or low-ability students are assigned to small classes.

A number of research approaches have been used to overcome the problem of bias. The approach we chose is to employ multiple regression methods that use "instrumental variables." This requires an additional variable in the data set that affects class size but does not directly affect student learning. Caroline Hoxby used this

[1] Technically, the estimation problem arises when school quality characteristics that are not in the data influence parents' choices of schools for their children; if all school quality characteristics that influence parents' choices are in the data, standard multiple regression methods (explained below) would provide unbiased results.

approach in a study of public schools in Connecticut published in the November 2000 issue of *Quarterly Journal of Economics*. More specifically, she used an instrumental variable approach to exploit variation in class size due to year-to-year variation in births in each school's "catchment area" (the geographic boundaries that determine the assignment of children to specific public schools). The basic idea of this method is simple. The number of children born in a school's catchment area varies each year due to random events that determine when children are born. This random variation in births generates random variation in class size when those children reach school age and enroll in their neighborhood schools. This random variation in class size over time in the same school can be used to estimate the impact of class size on learning in each school; these estimates are based on variation within schools in class size, not variation across schools in class size. This method requires at least two years of data for each school because estimates are based on changes in class size over time within each school. Hoxby found no effect of class size on test scores in mathematics and reading among fourth-grade and sixth-grade students in Connecticut.

Estimates Based on Minnesota Data

Almost all of the data used in our research are publicly available from the Minnesota Department of Education's Web site. The most important data are test score data. Beginning with the 1997–1998 school year, Minnesota has administered the Minnesota Comprehensive Assessment (MCA) test to third-grade and fifth-grade students each spring, usually in March.[2] The MCA test consists of math and reading for both third and fifth graders, plus a writing test for fifth graders only. Each year, about 60,000 third graders and 60,000 fifth graders in approximately 900 schools in Minnesota take the test.

Test score data are publicly available for the eight school years from 1997 to 2005, but some schools do not have test scores available for all eight years because they did not exist for all eight years, they did not participate in the test in some years, or they had fewer than 10 students in grades 3 or 5 in a given year (in which case data are not publicly available to protect students' privacy). Because comparisons of unadjusted test scores are not very informative, our research uses test scores divided by the standard deviations of the distribution of students' scores for the regression analysis, as is done in almost all research of this type.

The demographic data available from the Minnesota Department of Education consist of the number of children in each grade by race, gender, eligibility for subsidized lunch, limited English proficiency, and special education status. As will be seen below, all of these variables have strong predictive power for student test scores.

Unfortunately, not all of the elementary schools in Minnesota can be used in the regression analysis. About 1–2% of the schools were dropped from the sample because demographic variables used in the regression are not available. Another 8–9% of schools were not included because they have fewer than 10 students in grades 3 and 5, and so test scores are not reported for privacy reasons. Another 7–8% of the schools were dropped because they did not have test score data for at least two years, which is essential for doing estimation with school fixed effects.

Finally, data on class size are not available from the Minnesota Department of Education Web site; the site includes total enrollment for each year, but to obtain class size information, total enrollment must be divided by the number of classes (or the number of teachers). This information is often available from a school's Web site or by simply asking parents whose children go to a given school, but information for past years is more difficult to obtain. To obtain the number of classes in past years, we simply called the schools to ask. Many schools were able to provide this information, but others were not. Missing data on the number of classes caused about 35% of the schools to be excluded from our analysis.

Fortunately, comparisons of schools with and without information on the number of classes in past years reveal no significant differences in any of the other variables. Table 1 shows summary statistics on enrollment and school level means for demographic characteristics, both including and excluding the schools with missing data on the number of classes (and thus missing data on class size). Examining the larger sample, the average school had about 74 students in grade 3 and about 80 students in grade 5. About 7% of students in both grades were Black and 4% were Latino; about 82% were White. In each year, about one-third of children were eligible for a free or subsidized lunch.[3] Approximately 6% of students had limited English proficiency, and 12–13% were enrolled in special education classes.

The descriptive statistics for the smaller sample, for which class size can be calculated, are almost identical to those of the larger sample, in terms of both means and standard deviations. This suggests that missing data on the number of classes in each school is random, so there is little cause for concern about selection bias. The average class size for grades 3 and 5 in each school was calculated by dividing grade enrollment provided by the Minnesota Department of Education by the number of classes for those grades. In the smaller sample, the average school has a class size of 22.4 in grade 3 and 24.4 in grade 5.

As mentioned above, parents, teachers, and school administrators usually agree that reducing class size has a strong impact on learning. What evidence is there from Minnesota to support this claim? The simplest way to examine this is to group schools by small and large class sizes and examine their mean test scores. This is done in Table 2.

When Minnesota's elementary schools are divided into two groups of equal size—those with class sizes that are less than the median and those with class sizes greater than or equal to the median—there is very little difference in the test scores of grade 3 and grade 5 students across the two groups. Indeed, counter to conventional wisdom, the test scores are slightly higher in the schools with larger class sizes, for both grades 3 and 5 and for both reading and mathematics. When schools are divided into three groups of equal size (bottom half of Table 2), among grade 3 students the middle group has higher reading and math scores, whereas among grade 5 students the group with the highest class sizes has the highest reading and math test scores.

Yet these simple comparisons are not particularly convincing because schools with different class sizes could vary in other ways. Multiple regression analysis offers a simple method to control for

[2] Beginning in 2004, the MCA test was expanded to higher grades. In 2006, the MCA test was revised and is now referred to as MCA-II. We use data only up through the 2004–2005 school year, and only for grades 3 and 5.

[3] Children from families whose income is less than 185% of the federal poverty line are eligible for reduced-price lunches, and children from families with incomes less than 130% of the poverty line are eligible for free lunches.

Table 1. Descriptive Statistics, by Grade, for Schools with and without Class Data

	Schools with demographic data and ≥ 2 years of test data		Schools for which there is also data on the number of classes	
	Mean	Standard Deviation	Mean	Standard Deviation
Grade 3				
Enrollment	73.7	43.4	75.9	41.3
Class size	—	—	22.4	4.7
White (%)	81.5	24.2	81.3	23.9
Black (%)	6.9	14.4	7.2	14.0
Latino (%)	3.8	7.5	3.8	7.3
Asian (%)	4.9	9.7	5.0	8.8
American Indian (%)	2.3	7.8	2.1	6.9
Male (%)	51.2	7.5	51.1	7.2
Eligible for subsidized lunch (%)	33.8	22.2	32.3	22.6
Limited English proficiency (%)	6.0	11.8	6.1	11.1
Students in special education (%)	11.8	6.5	11.8	5.9
Number of schools	922		502	
Grade 5				
Enrollment	80.1	58.3	79.8	45.7
Class size	—	—	24.4	5.4
White (%)	81.7	24.0	81.1	24.4
Black (%)	6.9	13.9	7.4	14.5
Latino (%)	3.4	6.6	3.5	6.8
Asian (%)	5.1	10.0	5.3	9.8
American Indian (%)	2.3	7.6	2.2	6.4
Male (%)	51.3	7.3	51.2	7.2
Eligible for subsidized lunch (%)	33.1	22.1	32.1	22.9
Limited English proficiency (%)	5.4	10.8	5.6	10.6
Students in special education (%)	13.5	6.5	13.4	5.8
Number of schools	895		482	

Note: These averages are calculated for all years from 1997–1998 to 2004–2005 for which data were available on a particular variable. Each school year is given equal weight.

differences in other *observed* variables. *Ordinary least squares (OLS) multiple regression* methods estimate the (linear) relationship between a "dependent" variable, in this case test scores, and one or more "explanatory" variables. If there is only one explanatory variable, OLS amounts to drawing a line with the "best fit" for a set of data points in a two-dimensional scatter plot graph. This method can be extended to two or more explanatory variables. The estimated relationship can be expressed in an equation where the dependent variable, call it "*t*" (for test score), is predicted by the sum of several explanatory variables, call them x_1, x_2, etc. The impact of each x variable on t is measured by its coefficient, which is denoted by b. For example, when there are three x variables the equation is:

$$t = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + e,$$

where e (the "error") is the difference between the value of t predicted by the equation and the actual value of t. OLS multiple regression produces estimates of b_0, b_1, b_2, and b_3, and if the estimation method does not suffer from any problems of bias (a very big "if"), these coefficients can be interpreted as measuring the causal impact of the x variables on t. Of particular interest is the x variable that measures class size; the associated b estimates the impact of class size on student test scores.

OLS estimates of the determinants of student test scores are presented in Table 3. These estimates have negative coefficients on class size, which suggests that increased class size reduces student performance in both reading and mathematics in both grades 3 and 5. Almost all impacts are statistically significant at the 5% level (the one exception is significant at the 10% level)[4], and those combining across both grades into a single regression with a higher sample size yield coefficients on class size that are statistically significant at the 1% level. The coefficients on class size in Table 3, which range from -0.0040 to -0.0054, suggest that reducing class size by 10 students would increase test scores by about 0.04 or 0.05 standard deviations (of the distribution of student test scores), which translates to roughly 8 to 10 points on the test. Although the direction of these impacts is intuitively plausible (unlike the results in Table 2), their size is quite small. Assuming that test scores follow a normal distribution (which is a reasonable approximation for these data), a student who is in the middle of the distribution (at the 50th percentile), will move only up to the 52nd percentile if his test score increases by 0.04 or 0.05 standard deviations. This is a small benefit for a decrease in class size by 10 students (which would be quite expensive to implement). Yet, as explained above, there are reasons to think that simple OLS results may be biased.

Before turning to results from other estimation methods, note that the coefficients on the other variables that are shown in Table 3 have intuitively plausible effects. Relative to White students,[5] Black students (and to a lesser extent, American Indian students) perform worse in both subjects in both grades, whereas Asian students perform better (especially in math). The performance of Latino students is not significantly different from the performance of White students after one accounts for differences in income (as measured by eligibility for a subsidized lunch) and proficiency in English. Male students have lower reading scores than female students in both grades. Low income, as indicated by eligibility for a free or subsidized school lunch, is associated with lower test scores, as is limited English proficiency (which presumably affects many Latino students), and greater numbers of students in special education is also associated with lower scores.

The two pooled regressions in Table 3 are regressions that combine the data from grades 3 and 5 into a single regression. The impacts (that is, the coefficients) of all variables were allowed to differ by grade, but for almost all variables, the differences by grade were statistically insignificant. Thus combining the data in this way will not yield misleading results, and it has the advantage of estimating the impact of class size more precisely (although the problems of bias discussed

Table 2. Comparisons of Test Scores by Class Size

A. Dividing schools by class size into two groups of equal size

		Class size < median class size (2)	Class size ≥ median class size (1)
Grade 3	Average math test score	1,510.9	1,512.1
	Average reading test score	1,500.7	1,500.9
Grade 5	Average math test score	1,507.3	1,518.8
	Average reading test score	1,545.0	1,554.9

Note: Median class sizes are 22.8 and 25.0 for 3rd and 5th grades, respectively. Class sizes less than 10 or more than 40 were excluded from the calculation. The number of students tested was used to weight each observation.

B. Dividing schools by class size into three groups of equal size

		Bottom 1/3 (1)	Middle 1/3 (2)	Top 1/3 (3)
Grade 3	Average math test score	1,509.2	1,517.5	1,507.7
	Average reading test score	1,497.0	1,508.6	1,495.9
Grade 5	Average math test score	1,504.3	1,510.1	1,523.6
	Average reading test score	1,541.0	1,548.0	1,559.2

Note: The cutoff class sizes are 21.0 and 24.5 for 3rd grade, and 23.0 and 26.5 for 5th grade. Class sizes less than 10 or more than 40 were excluded from the calculation. The number of students tested was used to weight each observation. The number of observations for each column is 1,017 and 964 for 3rd and 5th grades, respectively.

[4] Statistical significance refers to the probability that the "true" *b* coefficient is zero. Statistical significance at the 5% level indicates that the estimated coefficient is sufficiently far from zero that the probability that the "true" value is zero is less than 5%. The analogous statement holds for statistical significance at the 10% or 1% levels.

[5] There is no coefficient for White students in Table 3 because they are the "base group," so that all race effects are measured with respect to that group. In effect, the coefficient on White students is set to zero. The same is true for female students. They are chosen as the base group (for comparison with male students), and thus the coefficient for female students is set to zero.

Table 3. OLS Estimates of Effect of Class Size on Math and Reading Test Scores

	Dependent variable = math test score			Dependent variable = reading test score		
	Pooled (1)	3rd grade (2)	5th grade (3)	Pooled (4)	3rd grade (5)	5th grade (6)
Class size	-0.0045 ***	-0.0054 **	-0.0040 *	-0.0040 ***	-0.0041 **	-0.0040 **
Black	-0.0050 ***	-0.0059 ***	-0.0038 ***	-0.0064 ***	-0.0067 ***	-0.0061 ***
Latino	-0.0008	-0.0017	0.0003	-0.0012	-0.0023	-0.0002
Asian	0.0052 ***	0.0045 **	0.0056 ***	0.0035 ***	0.0029 *	0.0038 **
American Indian	-0.0023 *	-0.0018	-0.0027 **	-0.0023 **	-0.0011	-0.0033 ***
Male	-0.0005	-0.0012	-0.0005	-0.0030 ***	-0.0036 ***	-0.0023 ***
Eligible for subsidized lunch	-0.0105 ***	-0.0072 ***	-0.0107 ***	-0.0102 ***	-0.0078 ***	-0.0102 ***
Limited English proficiency	-0.0051 ***	-0.0048 ***	-0.0053 ***	-0.0072 ***	-0.0064 ***	-0.0076 ***
Students in special education	-0.0073 ***	-0.0064 ***	-0.0087 ***	-0.0085 ***	-0.0082 ***	-0.0090 ***
R^2	0.48	0.40	0.54	0.60	0.56	0.64
Number of schools	526	464	433	527	464	438
Number of observations (an observation is one year for one school)	4,407	2,263	2,144	4,432	2,266	2,166

Note: The test score variable was normalized so that its mean is zero and its standard deviation equals one. These regressions also include dummy variables for each year (for any year, a dummy variable equals 1 for that year and 0 for all other years). Pooled regressions, which combine both grades, include grade dummy variables and include only schools for which observations were available for both grades.

*** Statistically significant at the 99% level of confidence, which means there is a less than 1% probability that the difference in scores is a result of chance.
** Statistically significant at the 95% level of confidence, which means there is a less than 5% probability that the difference in scores is a result of chance.
* Statistically significant at the 90% level of confidence, which means there is a less than 10% probability that the difference in scores is a result of chance.

above are not reduced by combining the data from both grades). The impacts of class size in the combined regressions are statistically significant at the 1% level. A final result, which is not shown in Table 3, is that the class size variable was interacted with the various demographic variables to see whether class size effects were particularly strong for certain types of students. For example, one might think that lower class size is particularly helpful to students from disadvantaged backgrounds. Yet it turns out that none of these interactions were statistically significant, which implies that reduced class size affects all students more or less equally.

The estimated impacts of these demographic variables are quite large compared to the impact of class size. Indeed, since these variables are measured as percents (i.e., they range from 0 to 100), multiplying their associated coefficients by 100 gives the impact on a single student of belonging to a particular group. For example, the coefficient of about 0.006 for Black students implies that, after controlling for other factors in the regression, Black students test scores are about 0.6 standard deviations below the scores of White students. Returning to the estimated impact of reducing class size, the estimated coefficients of 0.0040 to 0.0054 imply that reducing the average class size for Black students by 10 students will increase their test scores by only 0.04 or 0.05, less than 10% of the gap between White and Black students. The coefficients on eligibility for a subsidized lunch are even larger then those on Black students, ranging from 0.0072 to 0.0107, which implies that reducing class size by 10 students will close the gap between poor and non-poor students by only about 5%. Similarly, the estimates indicate that only about 10% of the gap between students with limited English proficiency and students who are more proficient in English can by removed by reducing class size by 10 students.

Yet OLS estimates are likely to be biased for several reasons, as explained above. This implies that one must turn to a method that avoids the bias that is caused by class size being correlated with unobserved factors that directly affect learning. This is in fact what Hoxby's method attempts to do by

Table 4. FE-IV Estimates of Effect of Class Size on Math and Reading Test Scores (instruments based on school-level estimates of enrollment trends)

	Dependent variable = math test score			Dependent variable = reading test score		
	Pooled (1)	3rd grade (2)	5th grade (3)	Pooled (4)	3rd grade (5)	5th grade (6)
Class size	-0.0047 **	-0.0028	-0.0066 **	-0.0040 **	-0.0058 **	-0.0024
Black	-0.0071 ***	-0.0074 ***	-0.0066 ***	-0.0075 ***	-0.0075 ***	-0.0073 ***
Latino	-0.0026	-0.0038 *	-0.0009	-0.0026 *	-0.0043 **	-0.0007
Asian	-0.0031 *	-0.0034	-0.0026	-0.0031 *	-0.0037	-0.0023
American Indian	-0.0057 ***	-0.0058 **	-0.0057 ***	-0.0065 ***	-0.0059 ***	-0.0071 ***
Male	0.0002	-0.0005	0.0009	-0.0017 ***	-0.0014 *	-0.0021 ***
Eligible for subsidized lunch	-0.0044 ***	-0.0045 ***	-0.0044 ***	-0.0042 ***	-0.0041 ***	-0.0042 ***
Limited English proficiency	0.0004	0.0009	-0.0005	-0.0043 ***	-0.0029	-0.0059 ***
Students in special education	-0.0083 ***	-0.0076 ***	-0.0090 ***	-0.0101 ***	-0.0101 ***	-0.0100 ***
R^2	0.10	0.09	0.12	0.17	0.16	0.17
Number of schools	526	464	433	527	464	438
Number of observations	4,407	2,263	2,144	4,432	2,266	2,166

Note: The test score variable was normalized so that its mean is zero and its standard deviation equals one. These regressions also include year-fixed effects. Pooled regressions include grade-fixed effects.

*** Statistically significant at the 99% level of confidence, which means there is a less than 1% probability that the difference in scores is a result of chance.
** Statistically significant at the 95% level of confidence, which means there is a less than 5% probability that the difference in scores is a result of chance.
* Statistically significant at the 90% level of confidence, which means there is a less than 10% probability that the difference in scores is a result of chance.

using “natural” variation in births over time within school catchment areas to predict variation in class sizes. The results from applying Hoxby’s method to the data from Minnesota are shown in Table 4. The estimated impact of class size on test scores is about the same as in Table 3, although slightly less precisely estimated (as indicated by the lower statistical significance). Of the four regressions that are grade-specific, the estimated negative impacts of class size on test scores are statistically significant at the 5% level for two regressions, grade 5 math and grade 3 reading. Both pooled regressions are statistically significant at the 5% level as well, and they show that reducing class size by 10 students would increase student test scores by about 0.04 to 0.05 standard deviations (of the distribution of student test scores). The results for the other variables in the analysis are very similar to what they were in Table 3. Yet even though these impacts of class size are statistically significant, they are still quite small, which confirms the finding in Table 3 that large reductions in class size are unlikely to have strong impacts on students’ academic achievement.

Conclusion

Our research on class size in Minnesota suggests that reducing class sizes in elementary schools in Minnesota would have only very small impacts on student learning as measured by test scores. This may be surprising to many parents, but it is consistent with recent research. Even among those studies that have found some correlation between class size and student learning, none predict that reductions in class size will result in large improvements in test scores—certainly not as large as the gaps in test scores found between White and Black students, between poor and non-poor students, and between students with and without limited English proficiency.

Because these reductions in class size are very expensive to achieve, our results imply that schools and parents need to look elsewhere for policies that can lead to sizeable increases in student learning. Exactly what such policies may be is an important topic for future research.

Photo © Will & Deni McIntyre/Corbis

The authors' research suggests that reducing class sizes in elementary schools in Minnesota would have only very small impacts on student learning as measured by test scores, and that schools and parents should therefore look elsewhere for policies that can improve student learning.

Although the method used here suggests that class size reductions would have small effects on student learning, it is important to point out that the data were fairly limited and, more generally, almost any non-experimental method could lead to biased results. The best research method is probably a randomized evaluation, as was done in Tennessee's Project STAR. Education authorities in Minnesota should seriously consider undertaking a similar study in Minnesota. Although randomized studies of education policies have been very rare in the United States (whereas, in contrast, randomized studies are quite common in health research), the U.S. Department of Education's Institute of Education Sciences has recently become much more supportive of these types of studies. Such research would not only provide a better assessment of the impact of class size, but more importantly it could be used to find more promising policies to improve the learning of all children in Minnesota's schools.

Paul Glewwe is a professor in the Department of Applied Economics at the University of Minnesota. His research focuses on education in developing countries, and more recently on education in the United States. He previously worked as a senior research economist at the World Bank. **Hyunkuk Cho** and **Melissa Whitler** are Ph.D. students in the Department of Applied Economics at the University of Minnesota.

The research presented here was supported by a Faculty Interactive Research Program grant from CURA. The program was created to encourage University faculty to carry out research projects that involve significant issues of public policy for the state and that include interaction with community groups, agencies, or organizations in Minnesota. These grants are available to regular faculty members at the University of Minnesota and are awarded annually on a competitive basis.

For a more detailed discussion of the research on which this article is based, see "Do Reductions in Class Size Raise Students' Test Scores? Evidence from Population Variation in Minnesota's Elementary Schools," by Hyunkuk Cho, Paul Glewwe, and Melissa Whitler, Department of Applied Economics, University of Minnesota. For a copy of the paper, contact Paul Glewwe by e-mail at pglewwe@umn.edu or by telephone at 612-625-0225.