

Class Size: What Research Says and What it Means for State Policy



Bruce Ayres

Grover J. "Russ" Whitehurst and
Matthew M. Chingos

EXECUTIVE SUMMARY

Class size is one of the small number of variables in American K-12 education that are both thought to influence student learning and are subject to legislative action. Legislative mandates on maximum class size have been very popular at the state level. In recent decades, at least 24 states have mandated or incentivized class-size reduction (CSR).

The current fiscal environment has forced states and districts to rethink their CSR policies given the high cost of maintaining small classes. For example, increasing the pupil/teacher ratio in the U.S. by one student would save at least \$12 billion per year in teacher salary costs alone, which is roughly equivalent to the outlays of Title I of the Elementary and Secondary Education Act, the federal government's largest single K-12 education program.

The substantial expenditures required to sustain smaller classes are justified by the belief that smaller classes increase student learning. We examine "what the research says" about whether class-size reduction has a positive impact on student learning and, if it does, by how much, for whom, and under what circumstances. Despite there being a large literature on class-size effects on academic achievement, only a few studies are of high enough quality and sufficiently relevant to be given credence as a basis for legislative action.

The most influential and credible study of CSR is the Student Teacher Achievement Ratio, or STAR, study which was conducted in Tennessee during the late 1980s. In this study, students and teachers were randomly assigned to a small class, with an average of 15 students, or a regular class, with an average of 22 students. This large reduction in class size (7 students, or 32 percent) was found to increase student achievement by an amount equivalent to about 3 additional months of schooling four years later.

Studies of class size in Texas and Israel also found benefits of smaller classes, although the gains associated with smaller classes were smaller in magnitude than those in the Tennessee STAR study. Other rigorous studies have found mixed effects in California and in other countries, and no effects in Florida and Connecticut.

Because the pool of credible studies is small and the individual studies differ in the setting, method, grades, and magnitude of class size variation that is studied, conclusions have to be tentative. But it appears that very large class-size reductions, on the order of magnitude of 7-10 fewer students per class, can have significant long-term effects on student achievement and other meaningful outcomes. These effects seem to be largest when introduced in the earliest grades, and for students from less advantaged family backgrounds.

When school finances are limited, the cost-benefit test any educational policy must pass is not "Does this policy have any positive effect?" but rather "Is this policy the most productive use of these educational dollars?" Assuming even the largest class-size effects, such as the STAR results, class-size mandates must still be considered in the context of alternative uses of tax dollars for



Grover J. Whitehurst is director of the Brown Center on American Education and a senior fellow in Governance Studies at the Brookings Institution.



Matthew M. Chingos is a fellow in the Brown Center on Education Policy at the Brookings Institution.

education. There is no research from the U.S. that directly compares CSR to specific alternative investments, but one careful analysis of several educational interventions found CSR to be the least cost effective of those studied.

The popularity of class-size reduction may make it difficult for policymakers to increase class size across the board in order to sustain other investments in education during a period of budget reductions. In that context, state policymakers should consider targeting CSR at students who have been shown to benefit the most: disadvantaged students in the early grades, or providing a certain amount of funding for CSR but leaving it up to local school leaders on how to distribute it.

In settings where state mandates on maximum class size are relaxed, policymakers need to bear in mind that the effect of any increase in class size will depend on how such an increase is implemented. For example, a one-student increase in the pupil/teacher ratio in the U.S. would reduce the teaching workforce by about 7 percent. If the teachers to be laid off were chosen in a way largely unrelated to their effectiveness, such as seniority-based layoffs, then the associated increase in class size might well have a negative effect on student achievement. But if schools choose the least effective teachers to let go, then the effect of increased teacher quality could make up for some or all of the possible negative impact of increasing class size.

State resources for education should always be carefully allocated, but the need to judiciously weigh costs and benefits is particularly salient in times of austere budgets. Class-size reduction has been shown to work for some students in some grades in some states and countries, but its impact has been found to be mixed or not discernable in other settings and circumstances that seem similar. It is very expensive. The costs and benefits of class-size mandates need to be carefully weighed against all of the alternatives when difficult decisions must be made.

A Context for Linking Research to Policy

There are a small number of variables in American K-12 education that are both thought to influence student learning and are subject to legislative action. Class size is one. Others include human resource policies, funding levels, curriculum, days/hours of instruction, and testing and accountability. Advocates for legislation on any of these topics are likely to appeal to research evidence as support for their position. That is appropriate and desirable as long as: a) the evidence is of high quality, b) it is relevant to the legislative action under consideration, c) conflicting evidence isn't ignored, and d) alternative courses of legislative action are similarly evaluated and compared.

The absence of any of these four conditions undermines the legitimacy of advocacy that is built on assertions about what “research says.” If the evidence is not of high quality it provides little or no support for any conclusions. For

instance, advocates for class-size reduction might cite evidence that students in smaller classes perform better on state examinations. But this simple correlation could be due to families with higher levels of education living in more affluent school districts that can afford smaller classes. Class size per se might have no more to do with student achievement than the condition of the schools' sports stadiums. "Evidence" that is this weak is no evidence at all.

Research can be of high quality but of questionable relevance to legislative action because the settings and circumstances of the research are so different from those at hand. For example, a number of well-designed studies of class size in the U.S. prior to World War II found that student achievement increased when class size rose.¹ But the nature of the population, the organization of schools, the characteristics of teachers and so many other things differ between now and the U.S. between the two world wars that the relevance of this research for current legislation is weak.

Considering the balance of the evidence is also very important. Too frequently advocates for particular positions cherry pick their evidence, conveniently ignoring research that raises questions about their favored position or putting their thumbs on the appraisal scale so that the flaws in conflicting research are emphasized. Advocates for and against class-size reduction have engaged in or been accused of engaging in such cherry picking for as long as there has been research on this issue and the prospect of legislation.²

Finally, and most importantly, all legislative action that requires appropriations involves choices. An appeal to evidence to support expenditures without consideration of the costs and benefits of all the options that are available can seriously mislead. With a limited and currently shrinking pool of state funds available to support K-12 education, the relative productivity of expenditures should be carefully considered. What are the costs and benefits of maintaining a cap on class size relative to other state-mandated uses of funds for education? And what are the costs and benefits of state mandates on specific uses of education funds relative to appropriations that allow more flexibility at the local level in how funds are spent?

Background on Class-size Reduction

Legislative mandates on maximum class size have been very popular at the state level. In recent decades, at least 24 states have mandated or incentivized

¹ Jonah Rockoff, "Field Experiments in Class Size from the Early Twentieth Century," *Journal of Economic Perspectives*, 23(4): 211–230 (2009).

² See, e.g. Eric A. Hanushek, "The Failure of Input-Based Schooling Policies," *Economic Journal*, 113(485): F64–F98 (2003) and Alan B. Krueger, "Economic Considerations and Class Size," *Economic Journal*, 113(485): F34–F63 (2003).

class-size limits in their public schools.³ Because the legislatively imposed limits have nearly always required a reduction in class size compared to the period prior to the legislation, these initiatives are called class-size reduction (CSR).

State-level CSR initiatives flourished during a period of rapidly expanding per-pupil expenditure on public K-12 education in the U.S. (per pupil revenue increased by 58 percent in real dollars in the last 20 years⁴). Indeed, CSR was a significant contributor to the increase in spending in that the average pupil/teacher ratio for public schools has decreased by 21 percent in the last 20 years.^{5,6}

The average U.S. pupil/teacher ratio in the public schools is currently 15.3.⁷ With an average U.S. teacher salary of approximately \$55,000,⁸ each student has an individual cost of about \$3,600 in teacher salary alone. With about 49.3 million public school students enrolled, a one-student decrease in class size from the present average would cost over \$12 billion a year in aggregate for the U.S.⁹ A one-student increase in class size would generate an equivalent savings. The costs of CSR are not limited to teacher salaries. More classrooms are needed for smaller classes. In our example of a one-student reduction in class size across the U.S., more than 225,000 additional classrooms would need to be added to the nation's stock. In any context \$12+ billion a year for any educational initiative is a large amount. By way of comparison, the federal government's largest single K-12 education program, Title I of the Elementary and Secondary Education Act, involves about the same level of annual expenditure as would a one-student reduction in the nation's average pupil/teacher ratio.¹⁰

With the end of federal stimulus funding and economic growth at low rates, 40 states are projecting shortfalls for their 2012 budget year. Some, including

³ Education Commission of the States, "State Class-Size Reduction Measures," Denver, Colorado: Education Commission of the States (2005).

⁴ http://nces.ed.gov/edfin/tables/tab_public_effort_show.asp?referrer=edfin

⁵ http://nces.ed.gov/programs/digest/d09/tables/dt09_064.asp

⁶ Note that the pupil/teacher ratio is nearly always smaller than class size because it includes teachers in specialized roles as well as regular classroom teachers. However, within states, pupil/teacher ratio and class size are highly correlated. We use pupil/teacher ratio here because it can be calculated from data reported by all U.S. school districts to the federal government, whereas class size cannot.

⁷ http://nces.ed.gov/programs/digest/d10/tables/dt10_068.asp

⁸ http://nces.ed.gov/programs/digest/d10/tables/dt10_083.asp

⁹ Presently there are 3.2 million teachers serving 49.3 million students in the public schools, which corresponds to 15.3 students for every teacher. Decreasing the pupil/teacher ratio to 14.3 would require hiring 226,000 additional teachers, which at \$55,000 per teacher would cost \$12.4 billion/year in salary costs alone.

¹⁰ These per-pupil cost estimates are very conservative compared to others in the literature that try to account for all costs of teachers, including fringe benefits and facilities. See, e.g., Douglas N. Harris, "Toward Policy-Relevant Benchmarks for Interpreting Effect Sizes: Combining Effects With Costs," *Educational Evaluation and Policy Analysis* 31(1): 3–29 (2009).

large states such as California, Texas, and Illinois, are projecting revenue shortfalls that are more than 20 percent of the size of the 2011 budgets.¹¹ For these states, there is no single solution. Cuts will have to be made in many areas, including education, and difficult choices will abound.

In this context, we believe it is useful to revisit research on the effects of class size on student learning, and to explore what the findings from that research have to contribute to the budget deliberations that many state legislatures are presently or will shortly be engaged in. Does class-size reduction have a positive impact on student learning? If so, by how much, for whom, and under what circumstances? What would be the likely effect of relaxing class-size mandates? What are the uncertainties in the conclusions that can be drawn from existing evidence about state CSR policies?

Research on Class Size

There is a large body of research on the relationship between class size and student learning. A 1979 systematic review of the literature identified 80 studies.¹² There are surely hundreds today. The vast majority of these studies simply examine the association between variation in class size and student achievement. The primary difficulty in interpreting this research is that schools with different class sizes likely differ in many other, difficult-to-observe ways. For example, more affluent schools are more likely to have the resources needed to provide smaller classes, which would create the illusion that smaller classes are better when in fact family characteristics were the real reason. Alternatively, a school that serves many students with behavior problems may find it easier to manage these students in smaller classes. A comparison of such schools to other schools might give the appearance that small classes produce less learning when in fact the behavior problems were the main factor.

The most credible studies of CSR have utilized either randomized experiments, in which students and teachers are randomly assigned to smaller or larger classes; natural experiments in which, for example, a sudden change in class size policy allows a before-and-after analysis of its effects; or sophisticated mathematical models for estimating effects that take advantage of longitudinal data on individual students, teachers, and schools. We limit our review to such studies.

Research that supports the effectiveness of smaller classes

The most influential and credible study of CSR is the Student Teacher Achievement Ratio, or STAR, study which was conducted in Tennessee during

¹¹ <http://www.usnews.com/news/articles/2011/01/14/10-states-with-the-largest-budget-shortfalls>

¹² Gene V. Glass and Mary Lee Smith. "Meta-Analysis of Research on Class Size and Achievement." *Educational Evaluation and Policy Analysis* 1(1): 2–16 (1979).

the late 1980s. Beginning with the entering kindergarteners in 1985, students and teachers were randomly assigned to a small class, with an average of 15 students, or a regular class, with an average of 22 students. Thus the reduction in class size (7 students, or 32 percent) was quite large. There are several research studies based on the STAR experiment. We examine two, including one that focuses on longer-term outcomes.

Krueger's analysis of the Tennessee STAR experiment finds that elementary school students randomly assigned to small classes outperformed their classmates who were assigned to regular classes by about 0.22 standard deviations after four years.¹³ This is equivalent to students in the smaller classes having received about 3 months more schooling than the students in the regular classes.¹⁴ This effect was concentrated in the first year that students participated in the program. In addition, the positive effects of class size were largest for black students, economically disadvantaged students, and boys.¹⁵ Krueger estimates that the economic returns to class-size reduction in Tennessee were greater than the costs, with an internal positive rate of return of about 6 percent.

A recent long-term follow-up of STAR participants into adulthood utilized IRS tax records to investigate a range of outcomes.¹⁶ The researchers find that students assigned to small classes at the beginning of elementary school are about 2 percentage points more likely to be enrolled in college at age 20. They did not find any evidence of an impact on incomes at age 27, but the income effects are measured with too much imprecision to warrant strong conclusions.

In summary, STAR researchers have found positive effects of early and very large class-size reductions on academic achievement in school and college attendance, with the economic benefits of the program outweighing the costs. These are important results from a very strong research design.

Rivkin, Hanushek, and Kain used a sophisticated statistical model to examine the effects of natural variation in class size in Texas in the mid-1990s.¹⁷ The study utilized longitudinal data from more than one-half million students in over three thousand schools. The researchers found positive effects of

¹³ Alan B. Krueger, "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics*, 115(2): 497–532 (1999).

¹⁴ The average student gains 0.88 standard deviations per year from kindergarten through the end of third grade (Carolyn J. Hill, Howard S. Bloom, Alison Rebeck Black, and Mark W. Lipsey, "Empirical Benchmarks for Interpreting Effect Sizes in Research," *Child Development Perspectives*, 2(3): 172–177 (2008)). The STAR effect size of 0.22 is 25 percent of that, which corresponds to 2.5 months of a 10-month school year.

¹⁵ Matthew M. Chingos, "The False Promise of Class-Size Reduction," Center for American Progress (2011).

¹⁶ Raj Chetty, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan, "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR," NBER Working Paper No. 16381, September 2010.

¹⁷ Steven G. Rivkin, Eric A. Hanushek, and John F. Kain, "Teachers, Schools, and Academic Achievement," *Econometrica*, 73(2): 417–458 (2005).

smaller class sizes on reading and mathematics in 4th grade, a smaller but still statistically significant effect in 5th grade, and little or no effects in later grades. Because the researchers used state assessment results that were only available beginning at 4th grade, they could not estimate class-size effects for the early grades that were studied in STAR. The estimated class-size effects for 4th and 5th graders in Texas were about half the size the K-3 effects in Tennessee.

International studies also provide positive evidence for the effects of class-size reduction. Angrist and Lavy took advantage of a class-size limit in Israel of 40 students. Whenever there are more students in a grade than 40 per teacher, a teacher and classroom must be added. The effect on class size in smaller schools can be dramatic. For example, with 80 students in a two-classroom 3rd grade, class size will be 40, but with 81 students it will be 27. The researchers find positive effects of smaller fourth- and fifth-grade classes, with effect sizes that are on the lower end of the range of those found in the STAR study.¹⁸ They do not find any effects on third-grade scores.

Studies with mixed results

In 1996, California enacted a K-3 CSR program designed to reduce class size by ten students per class, from 30 to 20, throughout the state. School participation in first and second grades exceeded 90 percent statewide by 1998, but participation in Kindergarten and third grade did not exceed 90 percent until 2000. This staggered introduction of CSR provided opportunities for researchers to study its effects. CSR created 25,000 new teaching positions in its first two years. Many of these positions were filled by teachers without certification or prior teaching experience. Other positions were filled by experienced teachers who switched grades or schools.

Jepsen and Rivkin carried out a sophisticated analysis to examine the influence of both the class-size reduction and the changes in the teacher workforce.¹⁹ They find positive effects for class-size reduction that are about half as large as those found in Tennessee. At the same time they find that increases in the numbers of new and not-fully-certified teachers offset much of these gains. In other words, students who ended up in the classrooms of teachers new to their classrooms and grades suffered academically from the teacher's inexperience by almost the same amount as they benefited from being in a smaller class. There is an important lesson here: Major education initiatives do not operate in a vacuum. Policies designed to affect one dimension of a student's educational experience are likely to affect others as

¹⁸ Joshua D. Angrist and Victor Lavy, "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics*, 114(2): 533–575 (1999).

¹⁹ Christopher Jepsen and Steven Rivkin, "Class Size Reduction and Student Achievement: The Potential Tradeoff between Teacher Quality and Class Size," *Journal of Human Resources*, 44(1): 223–250 (2009).

well. Other unintended negative consequences of California's CSR policy included an increase in class size in grades four and five²⁰ and the use of multi-grade classrooms.²¹

Woessman and West, taking advantage of differences in average class size between the 7th and 8th grades within schools, examined class-size effects on performance on international examinations in 11 countries around the world.²² They find educationally meaningful effects of smaller classes in a small number of countries, and a roughly even split between no effects and small effects in the remainder of the countries. Interestingly, the countries in which they find educationally meaningful positive effects of smaller classes are those with low salary levels for teachers and lower than average performance on international exams. A low average salary level for teachers suggests that a country is drawing its teaching population from a relatively low level of the overall capability distribution of all its employees. Thus, the countries studied by Woessman and West seem to have taken different paths, with some opting for relatively large numbers of poorly-paid teachers who perform better in smaller classes and others having relatively fewer but better-paid teachers whose performance isn't as affected by the number of students in class. In this regard it is worth noting that the East Asian nations that perform at higher levels than the U.S. on international exams have very large class sizes.

Dee and West used a nationally representative database of students to compare the outcomes of the same eighth-grade students who had attended different size classes in different subjects. They find no overall impact of class size on test scores, i.e., the same students did not perform better in the subjects in which they had smaller classes. There was, however, a small positive effect on test scores in urban schools, and modest overall positive effects on non-cognitive skills such as student attentiveness and attitudes about learning.²³

Studies with negative results

Arrayed against these positive and mixed findings for CSR are two credible studies that find no positive effects. Hoxby examined natural class size variation in Connecticut that was caused when natural population variation triggered a change in the number of classes in a grade in a school. For example, a small school that has 15 first-grade students in one year and 18 the next year would

²⁰ David Sims, "Crowding Peter to Educate Paul: Lessons from a Class Size Reduction Externality," *Economics of Education Review*, 28: 465–473 (2009).

²¹ David Sims, "A Strategic Response to Class Size Reduction: Combination Classes and Student Achievement in California," *Journal of Policy Analysis and Management*, 27(3): 457–478 (2008).

²² Ludger Woessmann and Martin West, "Class-Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in TIMSS," *European Economic Review*, 50(3): 695–736 (2006).

²³ Thomas S. Dee and Martin R. West, "The Non-Cognitive Returns to Class Size," *Education Evaluation and Policy Analysis* (forthcoming).

have a larger class during the second year. Additionally, a school that has set a class-size limit of 25 would have one second-grade class of 25 if there were 25 second-grade students but two classes of 13 if there were 26 students. Hoxby finds no relationship between class size and achievement in fourth and sixth grade (which should reflect class size in all previous grades). Hoxby does not even find class-size effects at schools that serve disproportionately large shares of disadvantaged or minority students.²⁴

A recent study by Chingos systematically examined the broad and expensive Florida CSR policy. In 2002, voters approved an amendment to the Florida state constitution that set limits on the number of students in core classes (such as math, English, and science) in the state's public schools. Beginning with the 2010-2011 school year, the maximum number of students in each core class would be: 18 students through grade 3; 22 students in grades 4 through 8; and 25 students in grades 9 through 12.

In 2003, the Florida Legislature enacted a law that implemented the amendment by first requiring, from 2003-04 to 2005-06, districts to reduce their average class sizes either to the maximum for each grade grouping or by at least two students per year until they reached the maximum. Beginning in 2006-07, compliance was measured at the school level, with schools facing the same rules for their average class size that districts faced previously. Beginning in 2010-11, compliance was measured at the classroom level.

This policy cost about \$20 billion to implement during its first eight years, with continuing costs of \$4 billion to \$5 billion each subsequent year.²⁵

Taking advantage of the staggered introduction of class-size reductions over time at the district and school level, Chingos utilized a sophisticated before-and-after analysis to examine the effects of the policy on student achievement between 2004 and 2009. He finds no evidence that the Florida policy had any impact on test scores in grades 3 through 8 (state-wide assessments in math and reading were not administered in the earlier grades).²⁶

Research summary

Despite there being a large literature on class-size effects on academic achievement, only a few studies are of high enough quality and sufficiently relevant to be given credence as a basis for legislative action. Because the pool

²⁴ Caroline M. Hoxby, "The Effects of Class Size on Student Achievement: New Evidence from Population Variation," *Quarterly Journal of Economics*, 115(4): 1239-1285 (2000).

²⁵ "2009-10 Florida Education Finance Program," DOE Information Database Workshop, Summer 2009, available at <http://www.fldoe.org/eias/databaseworkshop/ppt/fefp.ppt>.

²⁶ Matthew M. Chingos, "The Impact of a Universal Class-Size Reduction Policy: Evidence from Florida's Statewide Mandate," Harvard University, Program on Education Policy and Governance Working Paper. 10-03 (2010), available at http://www.hks.harvard.edu/pepg/PDF/Papers/PEPG10-03_Chingos.pdf

of credible studies is small; the individual studies differ in the setting, method, grades, and magnitude of class size variation that is studied; and no study is without issues, including those reviewed here, conclusions have to be tentative.

It appears that very large class-size reductions, on the order of magnitude of 7-10 fewer students per class, can have meaningful long-term effects on student achievement and perhaps on non-cognitive outcomes. The academic effects seem to be largest when introduced in the earliest grades, and for students from less advantaged family backgrounds. They may also be largest in classrooms of teachers who are less well prepared and effective in the classroom.

The Tennessee STAR experiment generates the largest estimate of the payoffs of a big decrease in class size. In Krueger's cost-benefit analysis, the return to the investment in smaller class sizes in Tennessee was slightly bigger than the costs of implementing the program. In other words, it paid its way.

All other studies of CSR generate either smaller estimates of the effects of variation in class size or find no effects at all. Getting a decent sense of the size of the effect that can be expected from reducing class size is obviously important to evaluating its benefits. Few voters would support a multi-billion dollar initiative that results in improvements in student outcomes (or any other desirable outcome, such as the population's health or vehicle gas mileage) that are too small to be noticeable.

One way to roughly estimate the size of class-size effects that is consistent with the existing literature would be to assume that the effects are linear, i.e., a reduction in class size by one student would generate 10 percent of the benefit of a reduction in class size by 10 students, and to assume that the effects diminish with each grade in school, with a reduction of a given number of students in 5th grade expected to have about half the effect of reduction of the same number of students in kindergarten.

The largest estimates of the magnitude of class-size effects are those produced by Krueger (1999), who found that the students in classes that were 7 to 8 students smaller on average than regular-sized classes performed about 0.22 standard deviations better on a standardized test. This means that students performed about 3 percent of a standard deviation better for every 1 student less in the class. These effects were generated largely by class-size reductions in kindergarten. If we take the effect by 5th grade to be half the size of the kindergarten effect, then a reduction in 1 student per class would generate approximately 1.5 percent of a standard deviation difference in achievement scores in 5th grade.

This means that on a statewide assessment such as the Texas Assessment of Knowledge and Skills (TAKS), which has a mean of about 700 and a standard deviation of about 100 at 5th grade for mathematics, a reduction in class size by one student would generate an improvement of 1.5 scale score points. Thus a statewide mean of 700 on TAKS would become a statewide mean of 701.5. Alternatively, an increase of class size by one student would lead to a statewide

mean of 698.5 on TAKS. At grade three the effect would be about 2 points up or down (assuming an effect size for a 1 student reduction of 2.0, which is 2/3rds of the effect for earlier grades in STAR). To put a one or two point change in student performance as a result of class size in context, the difference between the average scale scores of whites and blacks on TAKS at 5th grade is 65 points. Note that our estimates of a one to two point effect on TAKS of a one student change in class size are based on an upper bound for class-size effects based on Krueger's analysis. Estimates that averaged together effect sizes for all the studies we have reviewed, including the two that found no effects at all (Hoxby; Chingos), would obviously be considerably smaller.

Funding Class-size Reduction vs. Other Initiatives

When school finances are limited, the cost-benefit test any educational policy must pass is not "Does this policy have any positive effect?" but rather "Is this policy the most productive use of these educational dollars?" Assuming even the largest class-size effects, such as the STAR results, class-size mandates must still be considered in the context of alternative uses of tax dollars for education. Will a dollar spent on class-size reduction generate as much return as a dollar spent on: raising teacher salaries, implementing better curriculum, strengthening early childhood programs, providing more frequent assessment results to teachers to help guide instruction, investments in educational technology, etc.?

There is no research from the U.S. that directly compares CSR to specific alternative investments. In other words, the comparison condition for all CSR studies has been business as usual rather than, for example, a comparison of \$20 billion invested in smaller classes vs. \$20 billion invested in higher teacher salaries. Thus, estimates of effects and costs from different education investments have to be extrapolated and estimated from different studies, and this process is necessarily inexact. Nevertheless, Harris finds short-term rates of return for computer-aided instruction, cross-age tutoring, early childhood programs, and increases in instructional time that are all greater than those for CSR.²⁷ Whitehurst does not estimate costs, but finds effects on student achievement from choosing more effective curriculum; reconstituting the teacher workforce (for example by substituting Teach for America teachers for new teachers from traditional training routes); and enrolling students in popular charter schools in urban areas that are all as large or larger than those obtained from CSR.²⁸

The popularity of class-size reduction may make it politically difficult for

²⁷ Douglas N. Harris, "Toward Policy-Relevant Benchmarks for Interpreting Effect Sizes: Combining Effects With Costs," *Educational Evaluation and Policy Analysis* 31(1): 3–29 (2009).

²⁸ Grover J. Whitehurst. "Don't forget Curriculum," *Brown Center Letters on Education*, #3. October 2009, Washington, D.C.: The Brookings Institution.

policymakers to increase class size in order to sustain other investments in education, even in a time of budget austerity. In that context, state policymakers might consider targeting the reductions at students who have been shown to benefit the most: disadvantaged students in the early grades, or providing a certain amount of funding for CSR but leaving it up to local school leaders on how to distribute it. Much smaller classes for inexperienced teachers who need support in developing skills or for teachers who are responsible for struggling students may make more sense than across the board reductions.

The tradeoff between class size and teacher salaries needs to be very carefully considered. Effects on student achievement related to differences in teacher quality are very large. The same data from the Tennessee STAR study that demonstrates long-term effects for class-size reduction produces estimates of much larger effects for variation in teacher quality within schools. Thus, for example, while differences between large and small classes in early elementary school had no long-term effects on the earning power of adults, differences in classroom quality did.²⁹ With fixed or reduced state budgets to support K-12 education, maintaining class-size limits means a larger pool of teachers with lower salaries. It means that funds that might be devoted to raising teacher salaries across the board or selectively in hard to fill positions or for highly effective teachers will be limited. By one estimate, an increase in average class size by 5 students would result in an across the board increase of 34 percent in teacher salaries if all the savings were devoted to that purpose.³⁰ Higher salaries would likely draw more qualified people into the teaching profession, and keep them there.

In the current fiscal climate, it is clear that the yearly increases in funding in real dollars that have long been enjoyed by our nation's public schools are coming to an end for the foreseeable future. Many states and districts are contemplating cuts in funding that will require schools to make hard choices. So although the research literature has focused on the effect of reducing class size, the current policy debate concerns the other side of the coin—the consequences of increasing the size of classes. The potential for negative consequences of larger classes clearly needs to be weighed against the fallout from cutting other programs in order to preserve smaller classes—both academic programs and non-academic offerings such as athletics and the arts.

Another important point is that the effect of any increase in class size will depend on how such an increase is implemented. Our earlier rough calculation indicated that a one-student increase in the pupil/teacher ratio in the U.S., which would save over \$12 billion per year in salary costs alone, would decrease the

²⁹ Raj Chetty, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan, "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR," NBER Working Paper No. 16381, September 2010.

³⁰ Matthew M. Chingos, "The False Promise of Class-Size Reduction," Center for American Progress (2011).

teaching workforce by about 7 percent of the nation’s teachers. Many school districts and states across the nation are considering reductions in the teacher workforce on this order of magnitude. If the teachers to be laid off were chosen in a way largely unrelated to their effectiveness, such as “last in first out,” then the associated increase in class size could well have a negative effect on student achievement. But if schools choose the least effective teachers to let go, then the effect of increased teacher quality could make up for some or all of any negative effect of increasing class size.³¹

State resources for education should always be judiciously allocated, but the need to carefully weigh costs and benefits is particularly salient in times of austere budgets. Class-size reduction has been shown to work for some students in some grades in some states and countries, but its impact has been found to be mixed or not discernable in other settings and circumstances that seem similar. It is very expensive. The costs and benefits of class-size mandates need to be carefully weighed against all of the alternatives when difficult budget and program decisions must be made.

³¹ See, e.g., Donald J. Boyd, Hamilton Lankford, Susanna Loeb, and James H. Wyckoff, “Teacher Layoffs: An Empirical Illustration of Seniority vs. Measures of Effectiveness,” CALDER Brief 12 (July 2010) and Dan Goldhaber and Roddy Theobald, “Assessing the Determinants and Implications of Teacher Layoffs,” CALDER Working Paper 55 (December 2010).

About the Brown Center on Education Policy

Established in 1992, the Brown Center on Education Policy conducts research and provides policy recommendations on topics in American education. The Brown Center is part of The Brookings Institution, a private nonprofit organization devoted to independent research and innovative policy solutions. For more than 90 years, Brookings has analyzed current and emerging issues and produced new ideas that matter - for the nation and the world.

Brown Center on Education Policy

The Brookings Institution
1775 Massachusetts Ave. NW
Washington DC, 20036
202.797.6090
202.797.6144 (f)
<http://www.brookings.edu/brown.aspx>

Governance Studies

The Brookings Institution
1775 Massachusetts Ave., NW
Washington, DC 20036
Tel: 202.797.6090
Fax: 202.797.6144
www.brookings.edu/governance.aspx

Editor

Christine Jacobs

Production & Layout

John S Seo

Email your comments to gscments@brookings.edu

This paper is distributed in the expectation that it may elicit useful comments and is subject to subsequent revision. The views expressed in this piece are those of the authors and should not be attributed to the staff, officers or trustees of the Brookings Institution.