Panel 3: Considering effect sizes within the policy context

Penny Wise and Effect Size Foolish
Greg Duncan, Institute for Policy Research, Northwestern University

This session focused on the merits of incorporating cost and benefit analysis into effect size calculations for intervention studies. In the Tennessee Star class size experiment students were randomly assigned to classrooms that averaged 15 students or 22 students for an average of 2.3 years. Schanzenbach completed an analysis of the ACT test scores for many of the students in the Tennessee Star experiment. The estimate of a .15 standard deviation achievement impact on ACT tests scores, near the end of high school, is a decidedly small effect size according to Cohen's standards. A cost benefit perspective would examine what the .15 effect size is worth and what does it cost.

It is easy to translate the 22 versus 15 student class size into an approximate cost. The smaller class size cost an additional $11,500 per pupil. The cost-benefit question determines whether the dollar benefits exceed the $11,500 cost. One approach it is to try to translate test scores into productivity effects. The economics literature suggests that a one standard deviation increase in test scores produces an approximately 20 percent increase in lifetime earnings. Thus, a .15 standard deviation achievement impact translates into lifetime earnings gains of $17,000, some 1.5 times the cost of the intervention. Other possible benefits, such as crime reduction might also add to the benefit total. As a result, this benefit may well be a conservative estimate. The point to emphasize is that inexpensive programs with small effects may generate more benefits than costs and thus be worthwhile public investments. Conversely, expensive programs with big effects may cost more than they are "worth."

There are several factors when thinking about cost and benefit. First, rather than only considering the taxpayer perspective, the total social cost and benefits of a program also should be analyzed. This allows researchers to compute the costs to both taxpayers and participants. On the benefits side, researchers want to know not only to what extent taxpayers might benefit but also to what extent the participants themselves might benefit. It is the total taxpayer and participant perspective that matters because this represents the total resource cost of providing a particular intervention and the total gain to society of having people receiving it. If the total social benefits exceed the total social costs, than society as a whole benefits from the intervention.

Second, the outcomes that are measured are important in conducting the intervention. Researchers should consider a wide array of potential benefits—even broader than the intended program targets. It is especially important to measure outcomes that can be linked to important social costs such as grade failure, reduced crime, and higher productivity (earnings). For example, an early behavioral intervention might reduce grade failure or placement in special education and boost test scores. Using a broad array of outcomes also attempts to measure "spillover" benefits and costs. For example, does improving one child's behavior enable classmates to learn more? Improving child behavior may make a classroom more manageable for the teacher, resulting in a treatment benefit for other students in the classroom.

Third, not every outcome can be assigned a dollar amount. That is not to say that these types of outcomes are not important or that outcomes that can not be put in dollars should be ignored. The recommendation is to assign a dollar amount to as many outcomes as possible, but to the extent that there are important benefits. For example, even though tolerance and citizenship cannot be monetized researchers still measure their impact. After monetary benefits and costs have been tallied up, continue to assess whether these other benefits or costs might change conclusions. With this perspective, researchers employ discounting in which distant benefit dollars are worth much less than today's cost dollars.

Finally, not all interventions lend themselves to cost and benefit analysis, but there is a need for some attempt for order of magnitude estimates for costs and benefits. For most interventions, staff costs usually dominate because it is often hard to know the order of magnitude estimate for how many hours of professional time is actually spent per subject. In sum, by all means calculate effect sizes and translate effects sizes and costs into cost benefits for policy consideration.