World Development Vol. 77, pp. 262–276, 2016
0305-750X/© 2015 Published by Elsevier Ltd.
www.elsevier.com/locate/worlddev
http://dx.doi.org/10.1016/j.worlddev.2015.08.020

Cost-Effectiveness Analysis in Development: Accounting for Local Costs and Noisy Impacts

DAVID K. EVANS and ANNA POPOVA *
World Bank, USA

Summary. — Effective evidence-based policy making in development requires rigorous measurement of costs as well as impacts. This paper discusses important challenges and relevant solutions for implementing cost-effectiveness analysis and comparing the relative cost-effectiveness of programs across settings in the context of education. Adapting development programs from one context to another requires many assumptions. Most of the discussion of those assumptions, to date, has focused on the context-specificity (or external validity) of impact estimates. This study examines the sensitivity of cost-effectiveness analysis to errors in impact estimates, as well as the sensitivity of costs to context, and explores how biases such as recall and pilot bias may lead to over- or under-estimates of cost-effectiveness. We use data on the cost-effectiveness of 27 student learning programs and 16 attendance-boosting programs across Africa and Asia to demonstrate the magnitude of these challenges, as well as potential solutions. We show that comparing Monte Carlo simulations of cost-effectiveness to certain benchmarks and adapting the largest cost elements to local prices (i.e., parameter variation analysis) can remove much of the uncertainty surrounding cost-effectiveness estimates, and we propose that biases can be minimized through the use of detailed templates for cost reporting at the time of program implementation. Thus, this paper not only provides evidence of previously neglected challenges for cost-effectiveness analysis, but – more importantly – provides practical recommendations for undertaking this crucial stage in development planning well.

Key words — cost-effectiveness analysis, impact evaluation, education

1. INTRODUCTION

Recent decades have seen an increased emphasis on evidence-based decision making in development policy. As part of the trend, the past 20 years have witnessed a sharp rise in the implementation of rigorous impact evaluations of development programs.[1] Impact evaluations are a key tool for providing policy makers with evidence on what does and does not work to reduce poverty, expand investments in human capital, improve opportunities for women, and achieve other social objectives.

There is also an increasing awareness that analyzing impact alone is not enough to determine whether a particular program is worth investing in. The cost-effectiveness of the program or, in other words, how much it will cost to achieve a given impact, is also crucial. How this compares across programs should clearly affect evidence-based policy making. Consider, for example, a low-cost remedial tutoring program and a high-cost extension of the school day that both deliver the same improvement in test scores. Impact (or "effectiveness") alone would fail as a guide to policy, whereas cost-effectiveness would point to the lower cost program as the better investment. Cost-effectiveness analysis is proposed as an aid to compare the impacts and costs of various programs implemented in different countries and years when the programs have a common objective, making them comparable.

In brief, cost-effectiveness analysis calculates the incremental effect achieved by a given program per unit of its incremental cost. Incremental effects are the effects on a given outcome of interest for those receiving the program over and above those for a control group. These are often taken directly from the results of an impact evaluation and are thus expressed as an average treatment effect for a sample of individuals (e.g., average increase in school enrollment). Incremental costs reflect the additional resources used in a given intervention over the period for which it was evaluated, relative to those spent on a control group, measured in monetary-units (e.g., United States dollars per additional student treated).

While estimates of impact are reported directly in impact evaluations, cost can be calculated by following the "ingredients method," by which (1) all program resources (or "ingredients") are identified, (2) each ingredient is assigned a value (including its opportunity cost), (3) the values are then adjusted for inflation, time-value (since costs incurred in the future are worth less to society than those incurred in the present), and currency, and (4) the values are aggregated (McEwan, 2012). Once the costs have been aggregated, the cost-effectiveness ratio is calculated by simply dividing the effect of the program by its cost. Thus a program which increases student test scores by an average of 0.2 standard deviations with an incremental cost per student of $2 would have a cost-effectiveness ratio of 0.1 standard deviations per $1, for example.

Recent years have seen significant advances in tools to conduct cost-effectiveness analysis in development contexts, often applied to interventions in education. There is a growing literature documenting methods of cost-effectiveness analysis (see Levin & McEwan, 2001; Levin *et al.*, 2012; McEwan, 2012, and Glennerster & Takavarasha, 2013, for examples). At the same time, recent studies discuss some of the challenges in

* The authors are very grateful for comments and guidance from Amanda Beatty, Conner Brannen, Steven Glazerman, Rachel Glennerster, Kristen Himelein, Henry Levin, Patrick McEwan, Benjamin Olken, and three anonymous reviewers, as well as for feedback from participants in sessions at the Society for Research on Educational Effectiveness (SREE) Spring 2014 conference and the Center for the Study of African Economies Spring 2015 conference. The authors are especially grateful to Rachel Glennerster and the Abdul Latif Jameel Poverty Action Lab for sharing their cost-effectiveness data and analysis; and to Marie-Hélène Cloutier for providing additional data. Final revision accepted: August 18, 2015.

implementing cost-effectiveness analysis. These include, for example, the need for systematic methods (McEwan, 2012), the need for any reporting of cost data (McEwan, 2015), the sensitivity of cost-effectiveness estimates to errors in estimation of impact (Dhaliwal, Duflo, Glennerster, & Tulloch, 2013; McEwan, 2015), the context-specificity of impact estimates (Aronow & Samii, 2015; Meager, 2015), and choices around discount rates and exchange rates (Dhaliwal *et al.*, 2013).[2] Furthermore, some of the same studies have made advances in implementing a standardized approach to cost-effectiveness analysis. Dhaliwal *et al.* (2013) use randomized evaluations of 11 education programs from six countries to compare the cost-effectiveness for student attendance and enrollment, in terms of the additional years of student participation bought with $100 (in 2011 USD). They discuss the limitations of using simple point estimates of impact to conduct cost-effectiveness analysis, and carry out a sensitivity analysis which shows confidence intervals around cost-effectiveness estimates, using the standard error of impact estimates. Kremer, Brannen, and Glennerster (2013) likewise show cost-effectiveness across a range of learning interventions. These cost-effectiveness estimates are also reported and discussed in McEwan (2015).[3] In addition, researchers are undertaking both cost-effectiveness analyses (Lentz, Passarelli, & Barrett, 2013; Ryckembusch *et al.*, 2013) and cost–benefit analyses (Blunch, 2013; Glick, 2008; Whittington, Jeuland, Barker, & Yuen, 2012) of individual development programs.

The objective of this paper is to highlight certain neglected challenges in conducting cost-effectiveness analysis, including extrapolating the cost-effectiveness of programs across different settings, and to propose solutions to these challenges. We flesh out two key challenges for cost-effectiveness analysis – sensitivity to impact estimates and the context-specificity of costs – and propose solutions to these challenges. This study adds to the existing literature on sensitivity to impact estimates by proposing practical solutions for taking errors in impact estimates into account when using cost-effectiveness estimates for policy decisions. It then highlights a largely neglected challenge – how sensitive costs are to context – and demonstrates how costs can be adapted locally and when that is most likely to be successful. We also highlight several biases affecting cost-effectiveness analysis which have not been discussed extensively in the development literature, and put forward practical suggestions to minimize their effects.

In the next section, we demonstrate the sensitivity of relative cost-effectiveness estimates to errors in impact estimates and propose a systematic way to categorize the relative cost-effectiveness of programs in spite of this. Using J-PAL (2014) data, which compare the cost-effectiveness of 27 education programs in achieving student learning gains across Africa and Asia (14 of which have statistically significant impacts), and of 16 programs achieving attendance gains (11 of which have statistically significant impacts), we find that taking into account the 90% confidence interval around point estimates, we cannot rule out most rankings of the cost-effectiveness of programs with significant impacts. For policy makers choosing between competing programs based on their cost-effectiveness, this is a key finding. However, we propose using repeated simulations drawn from the distribution of the impact estimates (i.e., a Monte Carlo simulation) to translate this uncertainty into an intuitive measure of how often an intervention is likely to be at least as cost-effective as some pre-specified benchmark. Such a benchmark can either take the form of a particular intervention that has previously proven successful, or a sample of interventions among which a policy maker is choosing, in which case the output would be how often a given intervention is likely to fall toward the top of this distribution.

Subsequently, we examine questions of context-specificity (i.e., external validity) in cost measurement, which further complicate a simple approach to cost-effectiveness analysis, and propose parameter variation analysis as a solution.[4] We look at how costs vary across contexts and with program complexity, and how this complicates the extrapolation of cost-effectiveness results. We find that, using data on community teacher salaries in a number of countries, cost-effectiveness estimates vary by as much as 88% with the change of just one cost ingredient. However, varying the costs of parameters to which cost-effectiveness of a given program are most sensitive in this way can give us a much better approximation of how cost-effective an old program may be in a new setting. Such extrapolation is easiest for programs which either have few cost ingredients, or for which a large proportion of total cost is explained by few cost ingredients.

Finally, we explore how various biases – including "recall bias," much explored in relation to consumption but little in this literature, and "pilot bias," wherein pilot programs are likely to have higher costs but potentially also higher impacts – are likely to lead to biased cost-effectiveness estimates, and how these can be avoided by using templates to collect detailed information on costs at the time of implementation, as well as by distinguishing between fixed and variable costs when extrapolating cost-effectiveness ratios.

Cost-effectiveness analysis should be a crucial component of evidence-based policy making. However, for this analysis to be meaningful, governments and their advisors must carry out significant contextualization of cost estimates, build in uncertainty, and minimize biases before drawing conclusions about the cost-effectiveness of programs. We conclude with discussion and recommendations.

2. ERRORS IN THE ESTIMATION OF IMPACT

The issue of sensitivity to errors in the estimation of impact applies to impact comparisons in general, but its implications for cost-effectiveness analysis have been less discussed. Cost-effectiveness analysis typically uses point estimates of impact to calculate cost-effectiveness: For example, a girls' scholarship program in Kenya has an estimated impact of 0.27 standard deviations in student learning, with a cost per girl per year of $19.51, so the estimated cost-effectiveness is 1.38 standard deviations in student learning per $100 spent. Every point estimate, however, has a confidence interval, the size of which varies depending on the statistical power of the underlying evaluation. Imprecision in the estimation of impact means that it is possible that although one program may appear more cost-effective than a second when using point estimates to calculate cost-effectiveness, the relative cost-effectiveness may change – or the difference become trivial – if the variance around the two point estimates is taken into account.

We first demonstrate the extent to which this occurs in practice by re-ordering the relative cost-effectiveness for a sample of 14 education programs with student learning impacts, using the upper and lower bounds of the 90% confidence interval around the impact estimate and observing how this affects the relative ranking of the cost-effectiveness of these programs. Table 1 presents the original cost-effectiveness results of these 14 learning programs (Columns 1 and 2), alongside the results of this sensitivity analysis (Columns 6 and 7). We use data from J-PAL (2014), who conduct cost-effectiveness analyses of 27 education interventions across Africa and Asia, all of