

Comments on Impact Evaluation: Impact of Cost-Shared Water Supply Services on Household Welfare in Small Towns by Independent Evaluation Department, Asian Development Bank

Commentator: Emmanuel Jimenez, International Initiative for Impact Evaluation (3ie)

This report contributes substantively to an important development issue. It assesses the impact of the Asian Development Bank (ADB)'s Small Towns Water Supply and Sanitation Project (STWSSP), implemented during 2000-09 in Nepal. Because there has been very little evidence to date of the effectiveness of these types of water projects, which are community-based and have components to address financial sustainability, especially in small towns, the report should elicit a wide readership among development practitioners in the water sector. It is also clearly written – the report is easy to follow and readable, despite the technical details. The final version is attractively designed accompanying photos and layouts of graphs and tables.

I would recommend that the report be read, not only by sector specialists, but also by evaluators, who will find the its discussion on methodology of special interest. The report does a very good job in explaining why it was such a methodological challenge in conducting an evaluation that makes a convincing case for attribution. To their credit, the authors are very transparent with what they were up against. There was no baseline and there were no comparator groups identified at the beginning of the project. The team thus could only rely on comparisons of indicators based on ex-post surveys of project participants (“the treatment” group) and of a comparator group of non-participants (“the control” group). They made a valiant effort to reduce selection bias by using geo-spatial data, which are argued to be correlated with the original selection criteria, to select comparator control towns, as well as quasi-experimental matching methods of household data to select comparator households within those towns.

Impact evaluation specialists can argue whether or not this method has been fully successful and if the estimates of average treatment effects are precise. I'm sure many will. But this debate would be useful because what the team faced with this project is more likely to be the norm in most evaluations. The report's clear discussion of the choices the team had to make under the constraints would be valuable. Their challenges would make most readers agree wholeheartedly especially with the report's Recommendation 3, which is to “Strategically plan and implement impact evaluations for future programs and projects...” (p. xiv). By not considering evaluation at the design stage, the project missed a great opportunity to get more robust findings. Moreover this could have been done for relatively modest additional cost. As the paper points out, a proper baseline could have been done for all eligible towns. Such data would have been useful to select towns for inclusion. In addition, it is surprising that the team could not get access to “detailed town-level baseline data on selection criteria scores.” (p. xii). These were obviously computed for all towns that expressed interest in participating in the project. Had they been available, the evaluators could have used a regression discontinuity design that would arguably have been more robust than the matching procedure that the team had to follow.

Another reason this report would be of interest to evaluators is that, unlike many other impact evaluations, the paper uses its estimates of the effectiveness of the intervention to compare it to cost and re-estimates the rate of return. In order to make decisions about policy and programs, decision-makers need to know, not only about whether intervention works, but how much it costs relative to other interventions. While cost-effectiveness and cost-benefit analysis are built into evaluation guidelines, they are not all the time in impact evaluations. In education, according to David Evans of the World Bank, only about half of impact evaluations in education reported details on incremental costs and of those minimal details. You'd think economists would be at the forefront of such work, but they're not. David's blog which asks "[why don't economists do cost analysis in their impact evaluations?](#)", cites the fact that cost analysis isn't that interesting to economists or their journals because it feels more like accounting, not social science. It's really hard and time consuming to do and economists don't get much training on detailed cost-accounting. One can quibble about the assumptions made in the report but at least the benefit-cost analysis was done and, like the issue of methodology, can be the basis for a productive debate.

Finally, evaluation specialists would likely find of interest the paper's claim, that while there may be issues with internal validity, "The external validity of the design adopted for this evaluation can be considered better than randomized evaluation since it included an analysis along the causal chain of the TOC, the sample of households was representative of the population, the intervention has wide potential for application, and the economic logic of the findings applies more broadly." (p. 16) I thought that the statement needed further justification. After all, since the claims about this analysis in this quasi-experimental evaluation could also have been addressed with a randomized evaluation. Unlike the discussions on methods to control for selection and on costs, there was not as much detail on this issue that could fuel a useful debate.

In short, the report is overall a good read and will likely inform constructive discussions on sectoral issues as well as on tackling thorny evaluation questions like correcting for selection bias and comparing benefits to costs.