

Annex 1.14: Schooling Outcomes in Philippine Elementary Schools: Evaluation of the Impact of Four Experiments

I. Introduction

Project Description. In most developing countries high dropout rates and inadequate student learning in primary education are a matter of concern to policymakers. This is certainly the case in the Philippines: almost one-quarter of Philippine children drop out before completing sixth grade, and those who leave have often mastered less than half of what they have been taught. The government embarked on a Dropout Intervention Program (DIP) in 1990–92 to address these issues. Four experiments were undertaken: provision of multilevel learning materials (MLM), school lunches (SL), and each of these combined with a parent-teacher partnership (PTP). The first approach allows teachers to pace teaching to different student needs and is much less expensive than school feeding. Parent-teacher partnerships cost almost nothing but can help with student learning both at home and at school.

Highlights of Evaluation. The evaluation is noteworthy in that it explicitly aimed to build capacity in the host country so that evaluation would become an integral component of new initiatives, and data requirements would be considered before rather than after future project implementations. However, there are some problems that occur as a consequence, and the evaluation is very clear about what to expect. Another major contribution of the evaluation is the check for robustness of results with different econometric approaches. Finally, the benefit-cost analysis applied at the end is important in that it explicitly recognizes that significant results do not suffice: inexpensive interventions may still be better than expensive ones.

II. Research Questions and Evaluation Design

The key research question is the evaluation of the impact of four different interventions on dropping out and student outcomes. However, the evaluation design is conditioned by pragmatic as well as programmatic needs. The DIP team followed a three-stage school selection process:

- Two districts in each of five regions of the country were identified as a low-income municipality. In one district the treatment choices were

packaged as control, MLM, or MLM-PTP; in the other control, SL, or SL-PTP. The assignment of the two intervention packages was by a coin flip.

- In each district the team selected three schools that (a) had all grades of instruction, with one class per grade; (b) had a high dropout rate; and (c) had no school feeding program in place.
- The three schools in each district were assigned to control or one of the two interventions based on a random drawing.

Each intervention was randomly assigned to all classes in five schools, and both pre- and posttests were administered in both 1991 and 1992 to all classes in all 20 schools as well as in 10 control schools.

III. Data

The data collection procedure is instructive in and of itself. Baseline data collection began in 1990–91, and the interventions were implemented in 1991–92. Detailed information was gathered on 29 schools, on some 180 teachers, and on about 4,000 pupils in each of the two years. Although these questionnaires were very detailed, this turned out to be needless: only a small subset of the information was actually used, which suggests that part of the burden of the evaluation process could usefully be minimized. Pretests and posttests were also administered at the beginning and end of each school year in three subjects: mathematics, Filipino, and English.

The data were structured to be longitudinal on both pupils and schools. Unfortunately the identifiers on the students turned out not to be unique for pupils and schools between the two years. It is worth noting that this was not known a priori and only became obvious after six months of work uncovered internal inconsistencies. The recovery of the original identifiers from the Philippine Department of Education was not possible. Fortunately, the data could be rescued for first graders, which permitted some longitudinal analysis.

IV. Econometric Techniques

The structure of the sampling procedure raised some interesting econometric problems: one set for dropping out and one for test score outcomes. In each case there are two sets of obvious controls: one is the control group of schools, and the other is the baseline survey conducted in the year prior to the intervention. The authors handled these in different ways.

In the analysis of dropping out, it is natural to set up a difference-in-difference approach and compare the change in the mean dropout rate in

each intervention class between the two years with the change in the mean dropout rate for the control classes. However, two issues immediately arose. First, the results, although quite large in size, were only significant for the MLM intervention, possibly owing to small sample size issues. This is not uncommon with this type of procedure and likely to be endemic given the lack of funding for large-scale experiments in a developing-country context. Second, a brief check of whether student characteristics and outcomes were in fact the same across schools in the year prior to the interventions suggested that there were some significant differences in characteristics. These two factors led the authors to check the robustness of the results via logistic regression techniques that controlled for personal characteristics (PC) and family background (FB). The core result was unchanged. However, the regression technique did uncover an important indirect core cause of dropping out, which was poor academic performance. This naturally led to the second set of analysis, which focused on achievement.

A different set of econometric concerns was raised in the evaluation of the impact of the intervention INTER on the academic performance of individual I in school s at time t (AP_{ist}), which the authors model as

$$AP_{ist} = \delta_0 + \delta_1 AP_{ist-1} + \delta_2 PC_i + \delta_3 FB_i + \delta_4 LE_{st} + \delta_5 CC_i + \delta_6 INTER_{jt} + \varepsilon$$

where LE is learning environment and CC is classroom conditions.

First among these issues is accounting for the clustered correlation in errors that is likely to exist for students in the same classes and schools. Second is attempting to capture unobserved heterogeneity. And the third, related, issue is selection bias.

The first issue is dealt with by applying a Huber-White correction to the standard errors. The second could, in principle, be captured at the individual level by using the difference in test scores as an independent variable. However, the authors argue that this is inappropriate because it presupposes that the value of δ_1 is 1, which is not validated by tests. They therefore retain the lagged dependent variable specification, but this raises the next problem—one of endogenous regressor bias. This is handled by instrumenting the pretest score in each subject with the pretest scores in the other subjects. The authors note, however, that the reduction in bias comes at a cost—a reduction in efficiency—and hence report both least squares and instrumental variables results. The authors use both school and teacher fixed effects to control for unobserved heterogeneity in LE and CC .

The third problem is one that is also endemic to the literature and for which there is no fully accepted solution: selection bias. Clearly, because there are differential dropout rates, the individual academic performance

is conditional on the decision not to drop out. Although this problem has often been addressed by the two-stage Heckman procedure, there is a great deal of dissatisfaction with it for three reasons: its sensitivity to the assumption of the normal distribution, the choice and adequacy of the appropriate variables to use in the first stage, and its frequent reliance on identification through the nonlinearity of the first stage. Unfortunately, there is still no consensus about an appropriate alternative. One that has been proposed is by Krueger, who assigns to dropouts their pretest ranking and returns them to the regression. Thus the authors report three sets of results: the simple regression of outcomes against intervention, the Krueger approach, and the Heckman procedure.

V. Who Carried It Out

The data collection was carried out by the Bureau of Elementary Education of the Philippines Department of Education, Culture, and Sports. The analysis was carried out by a World Bank employee and two academic researchers.

VI. Results

The study evaluates the impact of these interventions on dropping out in grades one through six and on test score outcomes in first grade using a difference-in-differences approach, instrumental variable techniques, and the Heckman selection method. The effect of multilevel materials—particularly with a parent-teacher partnership—on dropping out and improving academic performance is robust to different specifications as well as being quite cost-effective. The effect of school lunches was, in general, weak. An interesting component of the study was a cost-benefit analysis—which makes the important point that the story does not end with significant results! In particular, a straightforward calculation of both the direct and indirect (opportunity) costs of the program leads to the conclusion that the MLM approach is both effective and cost-effective.

The lack of effectiveness of school feeding might be overstated, however: it is possible that a more targeted approach for school feeding programs might be appropriate. Furthermore, because there is quite a short period of time between the implementation and the evaluation of the program, the evaluation cannot address the long-term impact of the interventions.

VII. Lessons Learned

Several lessons were learned through this evaluation procedure. One major one was that the devil is in the details—that a lot of vital longitu-

dinal information can be lost if adequate information, such as the uniqueness of identifiers over time, is lost. A second one is that very little of the information that is gathered in detailed surveys was used and that a substantial burden to the respondents could have been reduced. Third, the study highlights the value of different econometric approaches and the advantages of finding consistency across techniques. Fourth, this study is exemplary in its use of cost-benefit analysis—both identifying and valuing the costs of the different interventions. Finally, although errors were clearly made during the study, the authors note that a prime motive for the study was to build evaluation capacity in the Philippines. The fact that the DIP was implemented and evaluated means that such capacity can be nurtured within ministries of education.

VIII. Source

Tan, J. P., J. Lane, and G. Lassibille. 1999. "Schooling Outcomes in Philippine Elementary Schools: Evaluation of the Impact of Four Experiments." *World Bank Economic Review*, September.