Impacts of an Early Education Intervention on Students' Learning Achievement

Evidence from the Philippines[1]

Futoshi Yamauchi[2]

Yanyan Liu[3]

International Food Policy Research Institute

Washington D.C.

June 2011

[1] We would like to thank seminar participants at the Philippine Department of Education, December 10, 2010 for their useful comments and the Japan International Cooperation Agency for financial support. We are most grateful to Yolanda Quijano for generous support and guidance from the onset of this project, and the Bureau of Elementary Education and various divisions within the Department for collaborations throughout this project including providing us with various databases for this study. Special thanks are offered to Juliet Abunyawan and Felisberta Sanchez who visited former TEEP division offices to collect TEEP investment data in addition to reorganizing DEDP database, and Ishidra Abunggol at the Research and Statistics Division who provided technical guidance to the first author. We thank Surajit Baruah for his excellent research assistance in managing the Basic Education Information System database. Any remaining errors are ours.

[2] International Food Policy Research Institute, 2033 K Street, NW, Washington D.C.; Email: f.yamauchi@cgiar.org

[3] Email: y.liu@cgiar.org

Abstract

This paper examines the impact of a large supply-side education intervention in the Philippines, the Third Elementary Education Project, on students' national achievement test score. First, the propensity score matching estimation, with local income level as the main conditioning factor, shows that the program was allocated to schools with low potential of growth and, both physical and soft components combined, significantly increased students' test scores at grade 4 to 6. The estimate is equivalent to about 15 score point increase with the six-year exposure. Interestingly, mathematics score is more positively responsive to the education reform than other subjects. Second, we found that textbooks, instructional training of teachers, and new classroom constructions particularly contributed to the above outcomes. Third, the empirical results also imply that early stage investments improve students' performance at later stages in the elementary school cycle, which suggests that public investments in the elementary education likely have dynamic impacts on education performance at the subsequent stages, increasing social returns to such an investment.

1. Introduction

It has been increasingly recognized that early stage investments are a critical input in human capital production. Early stage investments in the formation of human capital have dynamic impacts on human capital outcomes at the subsequent stages. Recent literature demonstrates that prenatal and early childhood nutrition status significantly determines the child's readiness for schooling and educational and labor market outcomes (Alderman, et al., 2000, 2006; Maluccio, et al. 2009; Yamauchi, 2008). The dynamic dependency of human capital formation on early stage investments essentially arises from the cumulative nature of its formation (Cunha, et al. 2006).

School education is not an exception. For instance, children cannot perform well at higher grades without sufficient acquisition of knowledge at lower grades. High repetition rates often observed at early grades in elementary school education indirectly proves the importance of initial investments in determining higher grade performance and affects returns to schooling (Behrman and Deolalikar, 1991). Similarly, successful completion at the elementary school stage determines performance at the secondary school stage though eventual catch up is feasible under some conditions.

In this paper, we assess the impact of a large scale intervention to elementary schools on students' learning performance in the Philippines. The Third Elementary Education Project (TEEP), implemented by the Philippine Department of Education (Dep Ed) in the period of 2000 to 2006 with financial assistance from the Japan Bank for International Cooperation and the World Bank, introduced a package of investments, both physical and soft components, to schools in the most depressed 23 provinces. It includes (i) school building construction and renovation, (ii) textbooks, (iii) teachers' training, (iv) school-based management (SBM), and (v) other facility and equipment support. The unique nature of TEEP was a combination of physical and soft components and institutional reform. While TEEP invests in physical buildings and textbooks, teachers and principals receive trainings and school and community are partnered to introduce school based management. Our study estimates the total impacts of the above investments and reform on students' learning performance, measured by their test scores.

Though we expect that such an intervention has dynamic effects on students' performance not only at the elementary school stage but at the secondary school and higher stages and even in labor markets,[4] the current study focuses on changes of test scores during the elementary school cycle. In this sense, our interest is in its short term impact. Moreover, the unit of observations in this study is a specific cohort at a school. Tracking one cohort of students (average) in each school, we aim to establish the TEEP impact on their test scores.

The second part of our analysis identifies component-wise impacts of school building construction and renovation, teachers' instructional and subject training, and grade-specific textbook distributions. Though the assessment of school-based management was also of our interest, we do not try to identify the effect since we found that the idea of SBM was basically introduced at an early stage (before our study period) and SBM fund allocation was endogenous (as a function of school's readiness). TEEP was initially designed with three batches at the province level, but these batch plans were altered in practice and the first and second batches in particular were implemented almost at the identical time.[5] In reality, therefore, we cannot use lags in implementation timings. Our supplemental data collection provides detailed school-level

[4] We collect individual and household data from 3500 students in 4 TEEP and 4 Non-TEEP divisions to study long-term impacts of TEEP. This component includes tracking of our sample students who migrated out of their origin communities.
[5] Khattri et al. (2010) used the lag between the first and second batches to identify the effect of school-based management on students' test score. Their analysis also includes TEEP investments such as new constructions as exogenous controlling variables. Their identification strategy is questionable given that, in reality, the initial phase plan was changed due to variations in preparedness across divisions.

information on implementation timings and investment amounts of different components, which also confirms that actual implementations did not follow the batch plan.

In our matching estimation, we use propensity score matching with municipality income class as the main conditioning variable. As described below, first, we observe that municipality-level income distributions are quite similar between TEEP and Non-TEEP divisions, which enables us to match TEEP and Non-TEEP schools from similar socio-economic backgrounds. In the Philippine context, local income level not only summarizes broadly socio-economic factors but also proxies the availability of private schools, which affects the competition between public and private schools and therefore the ability distribution of students in public schools (e.g., Yamauchi, 2005). Local labor market conditions are also controlled.

The paper is organized as follows. The next section discusses our identification strategy and Section 3 describes data used in our analysis. The empirical results are summarized in Section 4. We find that a two-year exposure to the TEEP intervention significantly increased test scores of grade-4 students. Our estimates show that test scores increased by 4.5 to 5 score points from grades 4 to 6, which implies about 15 score point increase if students are exposed to TEEP for 6 years of elementary school education (grades 1 to 6).

Component-wise impact analysis shows that (i) school building constructions and renovations, i.e., additional and renovated classrooms, (ii) instructional training of teachers, and (iii) additional textbooks significantly increased students' test scores. Interestingly, we found that investments in textbooks for earlier grades have dynamic effects on the performance of students at higher grades.

2. Identification strategy

We are interested in estimating 1) the average treatment effects (ATE) of TEEP on the performance of treated schools, and 2) how each component of TEEP contributes to school performance.

To estimate the ATE of TEEP on school performance, we combine double differences (DD) estimation with propensity score matching (PSM). PSM addresses the problem of purposive targeting of TEEP and DD helps wipe out cohort-specific fixed effects.

To illustrate our empirical approach, let $D=1$ if a cohort is treated (located in TEEP area) and $D=0$ if a cohort is not treated (located in a non-TEEP area). Let the outcome of being treated by TEEP and the counterfactual outcome at time t be denoted by (Y_t^T, Y_t^C). The gain from treatment is $(Y_t^T - Y_t^C)$ and we are interested in the average effect of treatment on the treated (ATET), $E(Y_t^T - Y_t^C|D_t = 1)$. As is known, the inability to observe the counterfactual outcome of treated households prevents us from estimating the ATT directly. Since we observe outcomes of the same cohort from 2002/03 and 2004/05, we use DD to control for cohort fixed effects. With $t = 1$ denoting 2004/05 and $t = 0$ 2002/03, we can write the standard DD estimator as

$$DD = E(Y_1^T - Y_0^C|D = 1) - E(Y_1^C - Y_0^C|D = 0) = E(Y_1^T - Y_1^C|D = 1) + B_1 - B_0,$$

where, B_t is the selection bias and $B_t = E(Y_t^C|D = 1) - E(Y_t^C|D = 0)$. If the selection bias is constant over time ($B_1 = B_0$), the DD estimator yields an unbiased estimate of the actual program impact.

The condition $B_1 = B_0$ or $E(Y_1^C - Y_0^C|D = 1) = E(Y_1^C - Y_0^C|D = 0)$ will not hold if the cohort characteristics or initial conditions affect subsequent changes of the outcome variables and have different distributions in the treatment and control groups. To account for this, we use PS matching to balance cohort characteristics and initial conditions. The assumption underlying PS matching is that, conditional on observables, X, the outcome change if not treated is independent of the actual treatment, i.e., $[(Y_1^C - Y_0^C) \perp D|X]$. This has been shown to imply $[(Y_1^C - Y_0^C) \perp D|P(X)]$ where $P(X)$ is the propensity score, defined as $P(X) = \Pr(D = 1 | X)$ (Rosenbaum and Rubin 1983).

We use a PS-matched kernel method and a PS-weighted regression method (Hirano, Imbens, and Ridder 2003). The PS-matched method estimates

$$\left[\sum_{D_i=1} (Y_i - \sum_{D_j=0} W_{ij} Y_j) \right] / N_1, \tag{1}$$

where N_1 is the number of treated villages, W_{ij} is the weight corresponding to villages i (treated) and j (untreated), and

$$W_{ij} = G[(P(X_j) - P(X_i)) / b_n] / \left[\sum_{D_k=0} G[(P(X_k) - P(X_i)) / b_n] \right], \tag{2}$$

where $G(.)$ is a kernel function and b_n is a bandwidth parameter. We use bootstrapping with 100 replications to estimate the standard errors for the PS-matched Kernel method. We choose the PS-matched kenel method instead of the more commonly used nearest-neighbor matching in order to obtain valid bootstrapped standard errors (Abadie and Imbens, 2006a and 2006b).

The PS-weighted method recovers an estimate of the ATET as the parameter β in a weighted least square regression of the form

$$\Delta Y_i = \alpha + \beta D_i + \varepsilon_i \tag{3}$$

where weights equal one for treated and $\hat{P}(X)/[1 - \hat{P}(X)]$ for non-treated observations. See Chen, Mu, and Ravallion (2009) for empirical applications of these two methods.

Since ATET can only be estimated consistently in the common support region of *X*, the choice of trimming method is important. We follow Crump *et al.* (2009) to determine the common support region by

$$A_{10} = \{X \mid P(X) \leq \lambda\} \tag{4}$$

where $\lambda = 1$ if

$$\sup_{X} \frac{1}{1-P(X)} \leq 2E\left[\frac{1}{1-P(X)} \mid D=1\right], \tag{5}$$

and otherwise solves

$$\frac{1}{1-\lambda} = 2E\left[\frac{1}{1-P(X)} \mid D=1, P(X) \leq \lambda\right] \tag{6}$$

This method minimizes the variance of the estimated ATET.

We use DD to do the component analysis which aims to estimate how each component of TEEP contributes to school performance. The empirical model is

$$H' = f(h,g,x; S) + H(S) + e$$

where H' and H are human capital stocks in SY 2004/05 and SY 2002/03 respectively (both are measured by test scores in SY2002/03 and SY2004/05 in our analysis), h is time and effort inputs (students), g is time and effort inputs (teachers) such as training, x is grade-specific physical inputs such as building and textbook, and S is the initial condition in school. By taking difference of human capital stock (test score) between the two years, we have

$$H' - H(S) = f(h,g,x; S) + e$$

We assume that the impacts of an intervention (i.e., change in physical and soft components) depends on the initial conditions in school, denoted by S, and students' effort.

Note that potentially H and (h,g,x) are complementary (thus, not separable) but we assume that the initial school conditions are sufficient to control such heterogeneities in the intervention effect. In other words, since TEEP was supposed to be targeted to poorly equipped schools and poor areas, it is important to control the initial conditions in the analysis. Trends could also differ between better-equipped (richer areas) and poorly-equipped schools (poorer areas). We use school district income classes and some school-level conditions.

Investment after SY2002/03 has to be adjusted with actual exposure time or differentiated by investment years. Our data identify when each component of TEEP was implemented at school level. Therefore, we know what investments were already made before SY 2002/03 and during SY2002/03 and SY 2004/05. For example, grade-4 students in SY 2002/04 can use grade-5 textbooks which were distributed in SY 2002/03. However, grade-6 students in SY 2004/05 could not use grade-4 and grade-5 textbooks which were distributed in SY 2004/05.

We estimate component-wise impacts in DD controlling the initial district- and school-level conditions (the conditioning variables being the interactions of school district income class and region)

3. Data

We use the Philippine school database provided from the Research and Statistics Division of the Philippine Department of Education. For our analysis, Basic Education Information System

(BEIS) data (SY2002/03 to SY 2008/09) and National Achievement Test (NAT) score data (SY002/03 to SY 2008/09) data are available. In NAT data, we have grade-4 test scores in SY 2002/03, grade-5 in SY 2003/04, and grade-6 in SY 2004/05 onward to the most recent year, SY 2008/09. Grade-4 in SY 2002/03 to Grade6 in SY 2004/05 is panel data tracking the same cohort in each school.

For TEEP implementation information, we have the Division Education Development Plan (DEDP) data, which was part of the TEEP completion reports. However, we found that the DEDP data do not identify implementation timings of different components and the completeness and quality substantially vary across divisions, though most of the worksheets contain highly useful information. The DEDP data has aggregated TEEP inputs over the period of SY 2000/01 to SY 2004/05.

To overcome the above data gap, we decided to visit 23 TEEP division offices to find the raw data on TEEP investments. The data cover details of different investments: textbook, training, school based management (SBM), school building (SBP), school innovation and improvement fund (SIIF), equipments/furniture, and supplementary instructional materials (SIM). Our team has encoded the data mostly from hard copies. Though some divisions had not recorded TEEP inputs as accurately as others, the data collection was successful.

For SBP, information on school constructions and renovations was received from the Dep Ed central office. However, since construction projects at the last stage of TEEP were not included in the database, we decided to complete the data by additionally gathering information on the last-stage projects from divisions.

For textbook and teachers' training, we used DEDP data formats to identify the implementation timings. For training, we only identified the starting date. For textbooks, we identified investment amounts (numbers and values by grade/subject) in each school year.

BEIS data provides very detailed information on student enrollment and achievements and teachers. The data normally disaggregate the information by grades, age, and gender. [6]. It is thus possible to characterize the initial school condition.

4. TEEP and Non-TEEP Divisions

For the short-term impact analysis, we decided to take following strategies. First, since TEEP was introduced to the most depressed provinces, the allocation of TEEP was purposeful. For example, TEEP was concentrated in CAR, a mountainous region in the center of northern Luzon. As Figure 1 shows, in Visayas, TEEP divisions were relatively scattered over space. In Mindanao, TEEP divisions were clustered, though it is not as clustered as those in northern Luzon. Under this circumstance, our sample use Visayas (regions 6, 7 and 8) to have relatively comparable TEEP and non-TEEP divisions in the same regions. It is important to choose relatively homogeneous areas to analyze TEEP impacts.

Figure 1 to be inserted

Second, we use the cohort panel from grade-4 (SY 2002/03) and grade-6 (SY 2004/05) to wipe out cohort-specific fixed unobservables. In this type of analysis, we have to consider two

[6] BEIS data needed an intensive programming to transform for analysis. The data was originally in Excel. To reorganize school-level data in different divisions/regions for one school year, we needed about 10 hours once our program ran. One year data has about 20 different sheets (each one of which contains huge data). We have completed this conversion for SY 2002/03 to SY 2006/07

sorts of unobserved fixed components: school-level and cohort-level. Since the unit of observations in the short-term impact analysis is school, we can always wipe out school-level fixed unobservables, but NAT data structure enables us to difference out cohort-specific effects too. If we pursue the long-term impact analysis using school data (not individual data from our tracking and household surveys), we can use grade-6 NAT score data in different years, handling only school-level unobserved fixed components.

We merged NAT grade-4 in SY 2002/03 and NAT grade-6 in SY 2004/05 using elementary schools in SY 2002/03. Therefore, we drop primary schools where only grades 1 to 4 students were taught. This restriction makes sense since our analysis used grades 4 to 6 in the cohort panel analysis. Therefore, our analysis pertains only to elementary schools in SY 2002/03 that offered grades 1 to 6. [7]

Note that TEEP also contributed to the conversion of primary schools to elementary schools by building new classrooms and staffing for grades 5 and 6. However, this effect is not included in our analysis. Moreover, it is possible that students from primary schools, not part of our sample, came into grades 5 and 6 in our sample elementary schools, which alters the cohort student compositions at grade 5.

Second, we needed income data on municipalities (or school district) to condition TEEP. The data we used came from Census 2000. Census 2000 defined income category (ranks 1 highest to 5 lowest) to each municipality. Note that some municipalities are split into a few school districts. For school districts in cities, we used rank 1 based on the income threshold used

[7] In this merging, we faced a technical problem: NAT SY 2002/03 (garde-4) data lacked new school IDs. The data recorded old IDs, since the Dep Ed introduced systematic new school IDs along with BEIS as part of TEEP. BEIS was started in SY 2002/03, but NAT SY 2003/03 data was not updated with the new school IDs. This created a difficulty in merging NAT SY 2002/03 and NAT SY 2004/05 (using the new school IDs). To prepare school IDs in SY2002/03 NAT data, we used BEIS SY 2002/03. They were merged with school names in each division (note that NAT data does not have district information). In this merging, our sample is restricted to elementary schools in SY 2002/03.

for municipalities. TEEP was not implemented randomly, but in divisions that were classified as those needing in the presidential social reform agenda.

Figures 2a, 2b and 2c to be inserted

Figure 2a shows the distribution of school districts by income class. School districts are concentrated in income classes 1, and 4 and 5, i.e., the highest income and lowest two income groups. In Figure 2b, we split the sample into TEEP and non-TEEP groups. Though we observe that more school districts are in income class 4 (less in income class 1) in TEEP group than non-TEEP group, the difference does not look significant. The distributions resemble between the two groups. Further, Figure 2c shows the distributions of schools in the two groups. Our basic observation remains valid here.

One important reason why local income levels need to be controlled is that the competition between public and private schools matters in the selection of students in the Philippine context. In high income municipalities (school districts), there are private schools that accept good students likely from well-off families. Therefore we expect differences in the ability distribution in public schools between high and low income municipalities. If school quality and student ability are complementary, the effect of TEEP on NAT change is expected to be different between high and low income districts.

Figures 2 confirm that there are income variations even within a TEEP division, so it is likely that we can find school districts that share similar socio-economic conditions in both TEEP and non-TEEP divisions in each region. This setting is quite helpful to our analysis for matching schools in similar income classes from TEEP and non-TEEP.

5. Results

5.1 Average treatment effect

The ATE of TEEP on school performance is estimated using the combination of DD and PSM. The first stage logit regression result is reported in Table 2. The dependent variable is 1 if the school is located in TEEP area and 0 otherwise. The explanatory variables include the interaction of municipality-level income categories and regional dummies, as well as school-level initial conditions (pupil-teacher ratio, grade 4 total enrollment, number of multi-grade classes, and proportion of local funded teachers). The results show that income categories, distinguished by regions, significantly explain TEEP implementation. Except income class 5, which is the poorest group, the effect is monotonic. In region 7 Central Visayas, which is omitted as the benchmark case, the effect of income class 5 is negative. In other regions, Western and Eastern Visayas, the income effect is monotonic over all income classes.

Tables 2 and 3, and Figure 3 to be inserted

The pseudo R-squared is 0.22, which suggests plausible explanatory power. The propensity score (PS) of each observation is estimated based on the regression. Figure 3 plots densities of the estimated PS in treatment and control groups as well as the cut-point of the PS values above which observations are trimmed. To illustrate the effects of trimming and reweighting, Table 3 displays simple differences of the explanatory variables between the treatment and control samples and the PS-weighted and kernel-matched trimmed samples. Although simple differences between the groups are large and statistically significant, trimming and matching based on the propensity score eliminates most significant differences except for one whose magnitude is considerably reduced.

Table 4 to be inserted

In Table 4, we report the estimation results on ATE of TEEP. We examine changes in overall and mathematics NAT scores from grade4 SY 2002/03 to grade 6 SY 2004/05. Panel 1 shows the simple DD results for the overall test and mathematics test scores. The effects on both scores are small in magnitude and insignificant statistically. Panels 2 and 3 show the results using DD + PSM (weighted regression) and DD + PSM (Kernel) respectively. The two methods give close results which suggest TEEP has significant impacts on both overall and mathematics scores. The magnitude is about 4 for the overall and 5 for mathematics. In other word, TEEP attributes to about 6% increase in the overall test score and 8% increase in the mathematics score on average.[8] The impact is not trivial over the two years' period. If the impact can continue at the same rate, the total effect of TEEP over 6 years (if students are exposed to TEEP in the entire elementary school period) would be about 12 to 15 score point increase. This magnitude of performance improvement is substantial. We note that the DD and PSM estimates of the TEEP impacts are larger than the simple DD estimates, which implies that the endogenous allocation of TEEP creates downward bias in the estimates if the program allocation is not taken into account. That is, it is likely that TEEP schools (and school districts) tended to have a lower trend

[8] This is computed by dividing the estimated ATET of TEEP by the counterfactual average score of the trimmed treatment group in SY2004/05.

in NAT than non-TEEP schools. For this reason, an estimate of TEEP effect using simple DD is smaller than the PSM estimates.

5.2 Component-wise results

In this analysis, we regress the change of overall test score over SY2002/03 and SY2004/05 on the various types of TEEP investment including textbook, training, and building, as well as school level initial conditions and the interaction of municipality-level income categories and regional dummies. As discussed, we examine the effects of textbook distributions, teachers' training, and school building constructions. The results are presented in Table 5, for the whole sample (Col. 1) and the TEEP sample only (Col. 2), respectively.

Table 5 to be inserted

There are some reservations. First, since our sample students (cohorts) are at grade 4 in SY 2002/03, we focus on textbooks for grades 4 to 6 distributed at TEEP. These students (cohorts) could have used TEEP textbooks at lower grades, but their impacts are reflected in their NAT scores at SY 2002/03 (grade 4). Second, teachers' training is split into two categories: theoretical and methodological training and subject-wise training. Their impacts could be different on students' performance. Third, though we have information on school building project contract values, we use the number of new constructions and renovations since the contract value aggregates both types and we also conjecture that the impacts are different between new constructions and renovations. The above conjectures were supported in preliminary analyses.

The findings are summarized as follows. First, in the textbook effect, earlier stage investments seem very important in determining later stage outcomes. Grade 4 textbook affects student outcomes from grade 4 to grade 6 onward. This finding is consistent with the recently well-established view on the cumulative process of human capital accumulation. Second, new classroom construction significantly helps improve their performance. The effect of renovations is insignificant. Third, conceptual/theory training seems to have a greater positive effect on student performance than subject-wise training (mathematics, English, and etc.). The latter has a negative effect on student performance at least in the short run for the whole sample probably because teachers have to use their teaching time for receiving training.

6. Conclusion

This paper provided evidence from the Philippines that both physical and soft components of public school education investments significantly increased students' test scores, by about 15 score points in national achievement test with the six year exposure. Our study also showed that the performance in mathematics is more positively responsive to the education reform and investments than other subjects.

Second, early stage investments improve students' performance at later stages in the elementary school cycle. The distribution of grade-4 textbooks is shown to increase their subsequent test scores more than grade-5 and/or grade-6 textbooks do.

The above findings, once combined with evidence in the literature, indicate that public investments in the elementary education likely have dynamic impacts on education performance at the subsequent stages, e.g., progression to high schools and colleges and academic performance therein though we do not analyze the effects beyond the elementary education stage. If so, social returns to such an investment can be greater than what the current study seems to show. This justifies large public investments to improve school quality available to students at the early stage of public education, whose cumulative benefits are realized at later stages in the education system and labor markets throughout their life cycles.

The competition between public and private schools is a unique feature of the Philippine education system due to the historical dominance of private institutions. In this context, some studies support ability screening hypothesis that private schools screen high-ability students but their actual schooling investments are not contributing to productivity increase (e.g., Yamauchi, 2005). The ability screening with the private-public competition, given high costs of private schools, is socially inefficient. If publicly subsidized and high quality education is available, we also expect the inflow of good students into the public school system in the long run.

References

Abadie, Alberto, and Guido W. Imbens, 2006a. "Large Sample Properties of Matching Estimators for Average Treatment Effects," *Econometrica*, 74: 235-267

----------, 2006b. "On the Failure of the Bootstrap for Matching Estimators," NBER Technical Working Papers 0325, National Bureau of Economic Research.

Alderman, Harold, Jere R. Behrman, Victor Lavy, and Rekha Menon, 2001. "Child Health and School Enrollment: A Longitudinal Analysis," *Journal of Human Resources*, 36: 185-205.

Alderman, Harold, John Hoddinott, and Bill Kinsey, 2006. "Long term consequences of early childhood malnutrition," *Oxford Economic Papers*, 58: 450-474

Behrman, Jere R and Deolalikar, Anil B, 1991. "School Repetition, Dropouts, and the Rates of Return to Schooling: The Case of Indonesia," *Oxford Bulletin of Economics and Statistics*, 53: 467-80

Chen, Shaohua, Mu, Ren, and Ravallion, Martin, 2009. "Are there lasting impacts of aid to poor areas?," *Journal of Public Economics*, 93: 512-528

Crump, Richard K., V. Joseph Hotz, Guido W. Imbens, and Oscar A. Mitnik, 2009. "Dealing with limited overlap in estimation of average treatment effects," *Biometrika*, 96: 187-199

Cunha, Flavio, James J. Heckman, Lance Lochner and Dimitriy V. Masterov, 2006, Interpreting the Evidence on Life Cycle Skill Formation, Erik A. Hanushek and F. Welch, eds. *Handbook of the Economics of Education,* Volume 1, Pages 697-812

Keisuke Hirano, Guido W. Imbens, and Geert Ridder, 2003. "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score," *Econometrica*, 71: 1161-1189

Khattri, Nidhi, Cristina Ling, and Shreyasi Jha , 2010, The Effects of School-Based Management in the Philippines: An Initial Assessment Using Administrative Data, Policy Research Working Paper No. 5248, World Bank

Maluccio, John A., John Hoddinott, Jere R. Behrman, Reynaldo Martorell, Agnes R. Quisumbing, and Aryeh D. Stein, 2009. "The Impact of Improving Nutrition During Early Childhood on Education among Guatemalan Adults," *Economic Journal*, 119: 734-763

Rosenbaum, P. R., and Rubin, D. B., 1983, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika* 70: 41–55.

Yamauchi, Futoshi, 2005. Why Do Schooling Returns Differ? Screening, Private Schools, and Labor Markets in the Philippines and Thailand, *Economic Development and Cultural Change*, 53: 959-81

Yamauchi, Futoshi, 2008. Early Childhood Nutrition, Schooling, and Sibling Inequality in a Dynamic Context: Evidence from South Africa, *Economic Development and Cultural Change*, 56: 657-682

Table 1: summary of the outcome variables

	TEEP				Non-TEEP			
	2002/03		2004/05		2002/03		2004/05	
Overall score	46.975	14.674	63.712	13.431	44.447	13.515	59.795	12.875
Math score	48.390	17.961	66.035	16.624	45.823	16.753	62.208	16.698
Number of obs.	1774		1774		2434		2434	

Table 2 Logit estimation of TEEP placement

Teep	Coef.	Std. Err.	Sig.
Region 6	-2.161	0.211	***
Region 8	-2.518	0.226	***
Income 2	1.341	0.308	***
Income 3	1.702	0.370	***
Income 4	0.306	0.190	
Income 5	0.141	0.186	
Region 6 * Income 2	-1.337	0.419	***
Region 6 * Income 3	-1.097	0.425	***
Region 6 * Income 4	0.330	0.259	
Region 6 * Income 5	-1.980	0.388	***
Region 8 * Income 2	-0.784	0.397	**
Region 8 * Income 3	-0.911	0.426	**
Region 8 * Income 4	1.325	0.264	***
Region 8 * Income 5	0.954	0.312	***
Pupil teacher ratio (both local and national)	-0.008	0.004	*
Grade 4 total enrollment (in ages 6 to 11)	-0.008	0.001	***
Number of multi-grade classes	-0.042	0.040	
Proportion of local funded teachers	0.203	0.596	
Constant	1.304	0.212	***
Number of obs.	4208		
LR chi2(18)	1258		
Prob > chi2	0		
Log likelihood	-2236		
Pseudo R2	0.22		

Table 3. Balance check

	diff1	se1	sig1	diff2	se2	sig2	diff3	se3
Region 6	-0.287	0.047	***	-0.004	0.046		-0.010	0.046
Region 8	-0.144	0.050	***	0.000	0.055		-0.003	0.057
Income 2	0.012	0.032		0.002	0.017		-0.004	0.022
Income 3	-0.012	0.040		0.000	0.035		-0.004	0.034
Income 4	0.108	0.050	**	0.004	0.062		0.022	0.060
Income 5	0.021	0.039		-0.001	0.054		0.000	0.041
Region 6 * Income 2	-0.024	0.015		0.000	0.010		-0.002	0.011
Region 6 * Income 3	-0.026	0.026		-0.001	0.025		-0.002	0.028
Region 6 * Income 4	-0.048	0.033		-0.002	0.032		0.001	0.038
Region 6 * Income 5	-0.101	0.020	***	0.000	0.005		-0.002	0.005
Region 8 * Income 2	-0.032	0.019	*	0.000	0.014		-0.004	0.014
Region 8 * Income 3	-0.041	0.027		0.000	0.025		-0.003	0.027
Region 8 * Income 4	0.026	0.038		0.001	0.047		0.003	0.044
Region 8 * Income 5	-0.008	0.014		-0.001	0.014		0.004	0.014
Pupil teacher ratio	-2.254	0.758	***	-1.101	0.847		-1.306	0.930
Grade 4 total enrollment	-7.475	1.325	***	0.687	1.198		0.511	1.257
Number of multi-grade classes	0.134	0.050	***	-0.037	0.077		-0.038	0.090
Proportion of local funded teachers	-0.005	0.003		-0.001	0.004		0.000	0.004
Number of observations	4208			3949			3949	

Table 4 Impacts of TEEP on school performance

	Untrimmed sample, simple DD				
	Treated diff	Control diff	**DD**	se	sig.
Overall score	16.737	15.348	**1.389**	0.874	
Math score	17.645	16.385	**1.260**	1.090	
Number of obs.	1774	2434			
	Trimmed sample, DD+PS weighted regression				
	Treated diff	Control diff	**DD**	se	sig.
Overall score	16.074	12.139	**3.934**	1.129	***
Math score	16.961	11.719	**5.242**	1.473	***
Number of obs.	1541	2408			
	Trimmed sample, DD+PS weighted kernel				
	Treated diff	Control diff	**DD**	se	sig.
Overall score	16.074	12.260	**3.813**	1.172	***
Math score	16.961	11.961	**5.000**	1.442	***
Number of obs.	1541	2408			

Table 5: Estimation results of component analysis

	TEEP & Non-TEEP			TEEP only		
	Coef.	Std. Err.		Coef.	Std. Err.	
Grade 4 textbooks (per pupil)	0.014	0.004	***	0.013	0.004	***
Grade 5 textbooks (per pupil)	-0.003	0.004		-0.005	0.004	
Grade 6 textbooks (per pupil)	-0.004	0.003		-0.002	0.003	
Instructional training (man-hours per pupil)	0.339	0.192	*	0.274	0.151	*
Subject training (man-hours per pupil)	-0.582	0.260	**	-0.331	0.236	
New constructions (SY2003/04)	2.287	1.199	*	3.364	1.099	**
New renovations (SY 2003/04)	0.235	0.292		0.547	0.319	*
Region 6	0.388	2.672		-5.387	3.186	*
Region 8	-2.808	2.716		-7.928	3.612	**
Income 2	5.629	2.906	*	2.924	3.025	
Income 3	-0.036	2.864		-1.827	2.854	
Income 4	-0.424	2.711		-0.940	2.855	
Income 5	1.666	2.526		0.992	2.667	
Region 6 * Income 2	-2.378	3.584		-4.206	3.840	
Region 6 * Income 3	-1.943	3.791		-1.516	3.557	
Region 6 * Income 4	-0.373	3.314		-5.089	4.279	
Region 6 * Income 5	0.467	3.156		-1.329	4.257	
Region 8 * Income 2	-1.671	3.738		-2.342	4.446	
Region 8 * Income 3	-0.382	3.349		3.524	4.528	
Region 8 * Income 4	0.066	3.187		0.943	4.178	
Region 8 * Income 5	2.788	3.473		2.117	4.180	
Pupil teacher ratio (both local and national)	-0.101	0.037	***	-0.095	0.051	*
Grade 4 total enrollment (in ages 6 to 11)	0.050	0.008	***	0.063	0.015	***
Number of multi-grade classes	-0.533	0.284	*	0.094	0.441	
Proportion of local funded teachers	-10.257	5.170	**	-11.512	9.816	
Constant	17.540	2.624	***	19.570	2.936	***
Number of obs	4186.000			1766.000		
F(25, 446)	5.870			8.230		
R-squared	0.046			0.108		

Figure 1 TEEP and Non-TEEP divisions in Visayas (Regions VI, VII, and VIII)

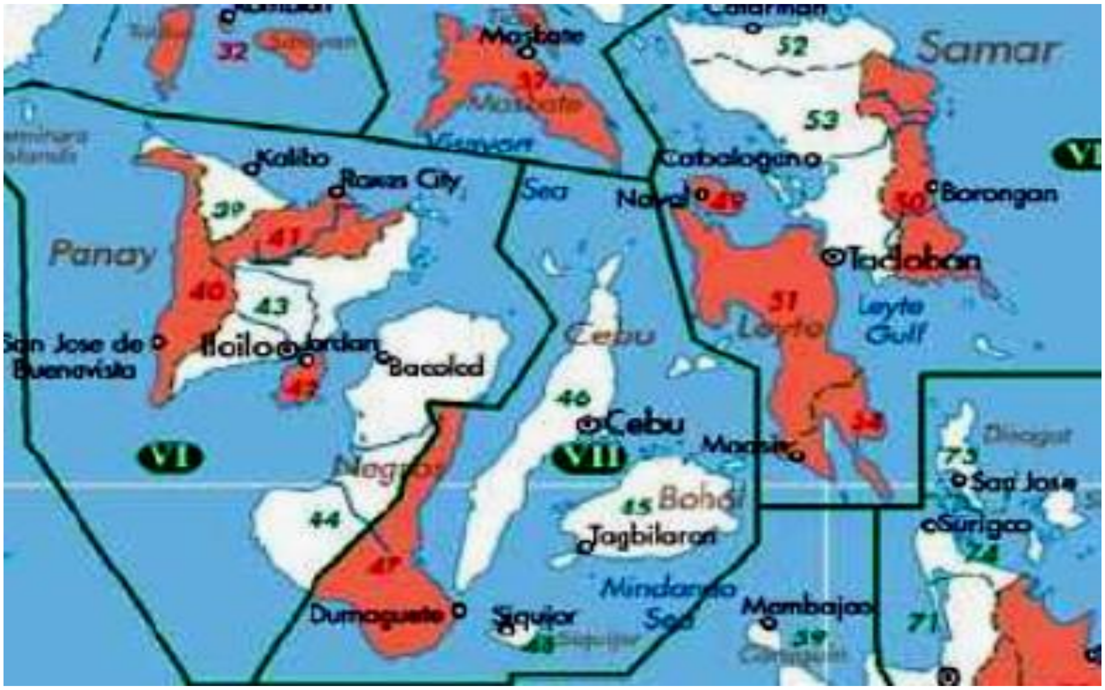

Figure 2a School District Income Class in Visayas (Census 2000)

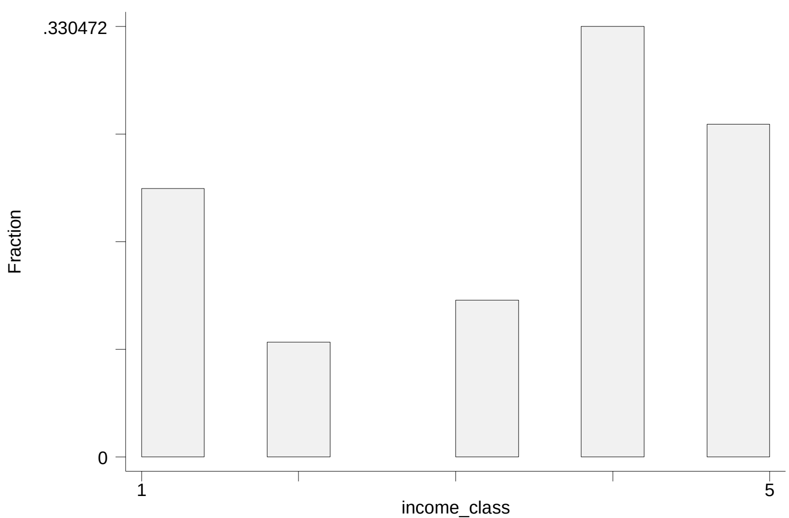

Figure 2b School District by Income Class (TEEP and Non-TEEP)

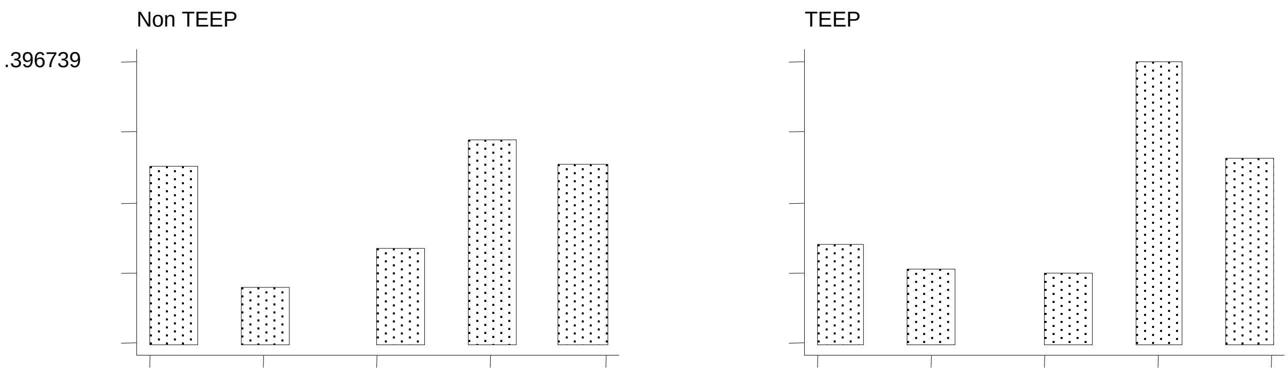

Figure 2c Schools by Income Class (TEEP and Non-TEEP)

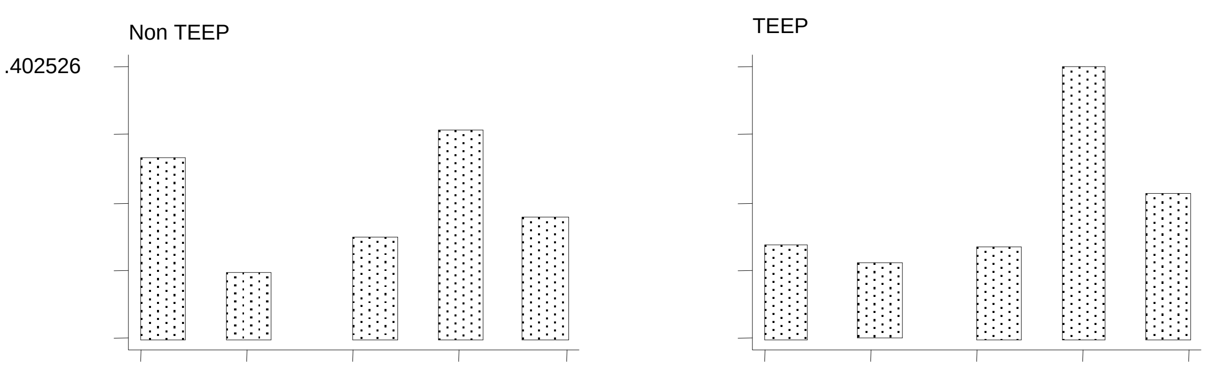

Figure 3 Plot of estimated propensity scores for Non-TEEP and TEEP districts