WPS6246

Policy Research Working Paper 6246

Impact Evaluation Series No. 72

Impacts of an Early Stage Education Intervention on Students' Learning Achievement

Evidence from the Philippines

Futoshi Yamauchi
Yanyan Liu

The World Bank
East Asia and the Pacific Region
Education Sector Unit
October 2012

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 scores. It finds that the program significantly increased student test scores at grades 4 to 6. The estimation indicates that two-year exposure to the program increases test scores by about 4.5 to 5 score points. Interestingly, the mathematics score is more responsive to the education reform than are other subjects. The analysis also finds that textbooks, instructional training of teachers, and new classroom construction particularly contributed to these outcomes. The empirical results imply that early-stage investment improves student performance at later stages in the elementary school cycle, which suggests that social returns to such an investment are greater than what the current study demonstrates.

This paper is a product of the Education Sector Unit, East Asia and the Pacific Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at fyamauchi@worldbank.org.

Impacts of an Early Stage Education Intervention on Students' Learning Achievement: Evidence from the Philippines

Futoshi Yamauchi[1]

Yanyan Liu[2]

Key Words: Human capital production, Education intervention, Philippines, Test Score

JEL Classification: I21, O15

[1] Correspondence: The World Bank, 1818 H Street, NW, Washington DC 20433; Email: fyamauchi@worldbank.org; Phone: 202-458-4262

[2] International Food Policy Research Institute, 2033 K Street, NW, Washington D.C

Acknowledgments

We would like to thank seminar participants at the Philippine Department of Education, the University of the Philippines, and JICA for their useful comments. 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 Third Elementary Education Project (TEEP) division offices to collect TEEP investment data in addition to reorganizing the Division Education Development Plan database, and Ishidra Afunggol 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. The TEEP student tracking survey conducted in eight provinces and the cities Manila, Cebu, and Baguio also offered enormous opportunities for the authors to extensively visit TEEP and non-TEEP schools and communities, which helped us correctly do our analysis in this paper. Any remaining errors are ours.

1. Introduction

Early-stage investments are increasingly recognized as a critical input in human capital production. These investments in the formation of human capital have dynamic impacts on outcomes at subsequent stages. Recent literature demonstrates that prenatal and early childhood nutrition status significantly determines a child's readiness for schooling and educational and labor market outcomes (Alderman et al. 2001; Alterman, Hoddinott, and Kinsey, 2006; Maluccio et al. 2009; Yamauchi 2008). The dynamic path of human capital formation depends on early-stage investments essentially due to 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. The high rates often observed of repeating early grades in elementary school show that many children face difficulty in successfully starting schooling, indirectly proving the importance of initial-stage investments in determining higher grade performance (Behrman and Deolalikar, 1991). Similarly, successful completion at the elementary school stage is a significant factor in student performance at the secondary school stage.

This paper assesses the impact of a large-scale intervention to elementary schools, the Third Elementary Education Project (TEEP), on students' learning performance in the Philippines. The project was implemented by the Philippine Department of Education from 2000 to 2006 with financial assistance from the Japan Bank for International Cooperation (JBIC; now JICA) and the World Bank. The unique nature of TEEP was in the combination of physical and soft components and institutional reform. Besides investing in physical buildings and textbooks, TEEP provided training to teachers and principals and introduced school-based management by partnering school with community. Our study estimates the total impacts of these investments and reforms on students' learning performance, measured by a change in student test scores during the elementary school cycle, though we expect that such an intervention has longer term effects beyond this stage, changing their activities in labor markets.[3]

Methodologically, we combine double differences with propensity score matching. We compare the change in test scores before and after the intervention in TEEP-treated schools with the change in nontreated schools. Propensity score matching is used to reduce the pre-intervention differences between the treated and nontreated schools. We find that a two-year exposure to the TEEP intervention significantly increased test scores in grade 4. Our estimates show that test scores increased by 4 to 5 score points (out of 100) from grades 4 to 6, which amounts to an increase of about 12–15 score points if students are exposed to the intervention for six years of elementary school education (grades 1 to 6). We also examine the effects of individual components of TEEP and find that school building constructions and renovations, instructional training of teachers, and additional textbook provision significantly increased student test scores. Interestingly, investments in textbooks for earlier grades have large positive effects on student performance at higher grades.

[3] We collect individual and household data from 3,500 students in four TEEP and four non-TEEP divisions to study long-term impacts of TEEP. This component includes tracking the sample students who migrated out of their original communities.

The paper is organized as follows: The next section describes the program. Sections 3 and 4 discuss data used in our analysis and our estimation method, respectively. Section 5 discloses the average treatment effects. The empirical results are summarized in Section 6. Section 7 concludes.

2. Program Background

The Third Elementary Education Project (TEEP) was implemented from 2000 to 2006 by the Philippine Department of Education in all public primary and elementary schools[4] in the 23 provinces[5] identified as the most socially depressed in the Social Reform Agenda.[6] The total project cost was US$221.16 million ($91.07 million from JBIC and $82.84 million from World Bank, $47.25 million from the Philippine government).[7] The unique feature of TEEP is a combination of investments in school facility and education materials and school governance reform. Not only were school facilities and textbook supply improved, but the decisionmaking process was also decentralized to the school and community levels. TEEP introduced a package of investments to schools in the selected 23 provinces. Specifically,

---

[4] Primary schools cover grades 1 to 4, while elementary schools cover grades 1 to 6.

[5] The program covered both primary (grades 1–4) and elementary (grades 1–6) schools. This paper analyzes the impacts on only elementary schools. However, converting primary schools to elementary schools by extending enrollment up to grade 6 was also an important part of the TEEP program. Students who complete primary schools are likely to attend elementary schools in grades 5 and 6, which changes the student body of those schools between grades 1–4 and grades 5 and 6.

[6] The Ramos administration, along with their medium term development plan, called *Philippines 2000*, identified reforms as the key to bridging social gaps and alleviating poverty. The objective of enhancing development through social reforms led to the formulation of the blueprint for social development in the Philippines, the Social Reform Agenda (SRA), marked as the first instance of social reforms in the history of the Philippines (Ramos 1995). As a result of the initial success of the SRA, the Congress of the Philippines in 1998 passed Republic Act 8425, widely known as the Social Reform and Poverty Alleviation Act (Republic of the Philippines, Congress, 1998). The law institutionalized the poverty alleviation program and a host of grassroots development strategies.

[7] See World Bank (2007, annex 1).

the package of investments included (1) school building construction and renovation, (2) textbooks, (3) teacher training, (4) school-based management, and (5) other facility and equipment support. Note that except principal-led projects on school building, schools or communities did not influence initiation of the above interventions.

The core of the program is school-based management, through which schools are given an incentive to manage proactively and more independently of the government. Schools were partnered with communities and parents to decide key issues such as improvement plan and school finance. Teachers were also trained systematically to improve teaching skills. Information management is being improved so that schools are responsible for systematically organizing information on enrollment, learning achievements, finance, and so forth and reporting it to the division office. Schools are required to set improvement plans every year and compare them with actual achievement. This dynamic process is monitored by the division-level education department. School finance is also being decentralized to some extent to relax the school budget constraints because Philippine public schools are not allowed to charge school fees. TEEP schools are free to raise their own funds from communities, parents, and others, though resources are admittedly limited in many poor communities. These reforms in public schools are expected to improve education quality, which would then in turn increase returns to schooling in labor markets (see Yamauchi 2005, on returns to schooling).

The selection of TEEP provinces was purposive because it intended to cover the most depressed provinces identified in the Social Reform Agenda. TEEP allocation is rather different in the Philippines' three macro-regions. As shown in Figure 2.1, in the northern macroregion of Luzon, TEEP was concentrated in the Cordillera Administrative Region, a mountainous region in the center of northern Luzon. In the central macro-region of Visayas, TEEP divisions were relatively evenly distributed. In the southern Mindanao macro-region, TEEP divisions were clustered, though not as clustered as in northern Luzon.

TEEP was initially designed to follow a phase-in plan with three batches at the province level. However, the plan was altered in practice due to variations in preparedness across divisions. Because understanding the implementation process of TEEP is important in choosing the appropriate strategy to identify the TEEP impacts, we collected school-level data on program implementation time and investment amounts of different components. Though the program implementation was substantially delayed,[8] it covered all schools in TEEP provinces. The data confirm that actual implementation did not follow the batch plan and suggest that the first and second batches were implemented almost simultaneously.[9] We will describe TEEP implementation in more detail in the data section.

[8] World Bank's TEEP Project Completion Report shows actual fund disbursements, which clearly proves substantial delays in the implementation (World Bank, 2007, page vii).

[9] Khattri, Ling, and Jha (2010) used the lag between the first and second batches to identify the effect of school-based management on student test scores. 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.

3. Data

This section describes the data used in our analysis. We combine the official test and school databases and the investment data that we collected in the (TEEP) divisions. For test scores and school conditions at the start of the project, we use the National Achievement Test (NAT) score data and the Basic Education Information System (BEIS) data, respectively. The NAT data provide average test scores for grade 4 students in school year (SY) 2002/03, grade 5 in SY 2003/04, and grade 6 in SY 2004/05 for each school. We note that grade 4 in SY 2002/03, grade 5 in SY 2003/04, and grade 6 in SY 2004/05 constitute panel data that tracked the same cohort in each school.[10] Double differences (DD) based on the cohort panel from grade 4 (SY 2002/03) and grade 6 (SY 2004/05) is used to eliminate cohort-specific fixed effects.[11]

Table 3.1 shows the mean and standard deviation of mathematics and overall scores of the cohort in SY 2002/03 and SY 2004/05 for TEEP and non-TEEP areas, separately. TEEP schools have significantly higher average scores than non-TEEP schools in both years. The BEIS data provide detailed information on student enrollment and achievements and teachers since SY 2002/03. The data normally disaggregate the information by grade, age, and gender.[12] Since BEIS was established as part of TEEP, we do not have systematically

[10] National achievement test is self-administered at schools, which potentially creates bias in raw test scores. For grade-6 students, tests were implemented in each year. In our analysis, however, we use an experimental introduction of the test for the same cohort: Grade 4 in SY 2002/03 and Grade 6 in SY 2004/05.
[11] Due to delayed preparations at the early stage of TEEP, most of the program schools received investments after SY 2002/03.
[12] BEIS data needed intensive programming to transform for analysis. The data were originally in Microsoft Excel. The computer program needed about 10 hours to reorganize school-level data in different divisions and regions for one school year.

collected well-managed school database before SY 2002/03, which makes it infeasible to examine the parallel trend assumption using the pre-intervention period.[13]

We obtain income data on municipalities (or school district) from the 2000 Census. Local income level is an important factor that determines school and family environments. Controlling local income levels is crucial because competition between public and private schools matters in the selection of students in the Philippine context. In high-income municipalities (school districts), students from well-off families and with high test scores are likely to be accepted into private schools. 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.

We assigned an income category to each school district based on the 2000 Census. The census defined income category (ranking from 1, highest, to 6, lowest) for each municipality.[14] Note that some municipalities are split into a few school districts. In cities, we ranked school districts as 1 based on the income threshold used for municipalities.

---

[13] We have also detected discontinuity of summary statistics of some key variables between before and after SY 1999/2001.

[14] The income classification of municipalities (municipality income) used in this paper is based on Republic of the Philippines, Department of Finance (2001), Department Order No. 32-01 (effective November 20, 2001) and Census 2000. The income categories for 1,435 municipalities are defined as follows: 1: Philippine peso (PHP) 35 million (M) or more (number of municipalities: 130); 2: PHP 27M or more but less than PHP 35M (140); 3: PHP 21M or more but less than PHP 27M (204); 4: PHP 13M or more but less than PHP 21M (543); 5: PHP 7M or more but less than PHP 13M (401); 6: less than PHP 7M (17). There are a small number of municipalities in income class 6 in the country, but the Visayas region (our sample) has income classes 1 to 5 only.

TEEP was implemented not randomly but in the divisions identified as socially most depressed in the presidential Social Reform Agenda. Figure 3.1 shows the distribution of school districts by income category in TEEP and non-TEEP groups. School districts are concentrated in income categories 1, 4, and 5—that is, the highest income and the two lowest income rankings—for both TEEP and non-TEEP. Though we observe that more school districts are in income category 4 (and fewer in 1) in the TEEP group than in the non-TEEP group, the difference does not look significant. Further, Figure 3.2 shows the distribution of schools in the TEEP and non-TEEP groups. Our basic observation remains valid here. Therefore, it is likely that we can find (and compare) school districts that share similar socioeconomic conditions in both TEEP and non-TEEP divisions.[15]

For TEEP implementation information, we have the Division Education Development Plan data, which was part of the TEEP completion reports. This dataset has aggregated TEEP inputs during SY 2000/01 to SY 2004/05. However, it does not identify implementation timing and inputs of different components of TEEP. Furthermore, the completeness and quality of the data substantially vary across divisions. To overcome this gap in the data, we visited 23 TEEP division offices to find the raw data on TEEP investments. The raw data we collected reveal details of different TEEP investments: textbooks, training, school-based management, school building, school innovation and improvement fund, equipment/furniture, and supplementary instructional materials. For training, we identified the starting date of teacher training and calculated the total number of man-

[15] In general, migration occurs from TEEP provinces since poverty rate is higher in those provinces. In the case of Visayas, we observe migrations to Manila and Cebu City (non-TEEP provinces which are not included in our analysis). Though the implementation of TEEP might have created an incentive for students to move from non-TEEP to TEEP schools, this is not realistic since many provinces are divided by mountains or sea (i.e., islands). See Figure 1.

hours spent in training during SY 2000/01 to SY 2004/05 by different categories. For textbooks, we identified investment amounts (quantity and cost by grade and subject) in each school year. Similarly, we sorted school building projects by completion year and identified new construction and renovation cases and their aggregate total values by school.

Table 3.2 describes the initial implementation timing of different TEEP components: school building new construction and renovation, textbooks, and teacher training. The table shows the percentage of schools covered under TEEP in Visayas (our analysis is restricted to this area) from SY 2000/01 through SY 2005/06. In school buildings, we aggregated new construction and renovation projects by their completion timings. In textbooks, we used timing in which textbooks (disaggregated by grade and subject) were distributed to schools. In teacher training, we only used the initial time when training was introduced at the school district. Note that training covers a wide range of contents, which principals and teachers studied step by step. In many cases, training was conducted at the school district level. This means that instructors visit districts one by one within a division, and therefore it took them a few years to cover all the topics (our data show only total man-hours and the start date). The table shows that by SY 2002/03, about 80 percent of schools had received textbooks and 50 percent had at least one completed school building project. In all schools, the training process had just begun.

4. Estimation Method

In this section, we describe our estimation methods. Because the allocation of TEEP was purposive, the initial school conditions are likely to have different distributions in the treatment and control groups. If the initial conditions affect subsequent changes of the outcome variables, DD would give a biased estimate of the TEEP impacts. We use two strategies to deal with the potential bias due to nonrandom program placement. First, we use the sample from Visayas only. As shown in Figure 2.1, TEEP divisions are relatively evenly distributed throughout Visayas compared with the other two macro-regions. We therefore expect that the TEEP and non-TEEP provinces are more comparable in Visayas, and hence our extra data collection and cleaning efforts were focused on Visayas. Second, we use propensity score (PS) matching to balance observable cohort characteristics and initial conditions between the treated and the control groups.

Because the original phase-in plan of TEEP was not strictly followed in the first two batches in practice, we cannot explore the pipeline design to identify the impact of TEEP on school performance. In this process, however, all schools in TEEP provinces were covered in the program. We also use all schools in non-TEEP provinces in the same region (regions 6, 7 and 8 of Visayas in our case). Therefore, we formed a control group based on all the schools in the non-TEEP provinces to estimate the counterfactual of the TEEP schools.

Three caveats exist in our method. First, our baseline is not free of contamination. Table 3.2 showed that, in Visayas region, TEEP had been at least introduced in many treatment schools by SY 2002/03. Thus, the initial level of test scores in the treatment group reflects earlier investments completed before SY 2002/03. Second, it is possible that students from primary schools, which are not part of our sample, came into grades 5 and 6 in our sample elementary schools, which alters the student body at grade 5. Since TEEP also contributed to the conversion of primary schools to elementary schools by building new classrooms and staffing for grades 5 and 6, it is possible that attrition is different in the treated and control groups.[16] Third, as an observational analysis, we cannot eliminate bias due to time-variant unobservables.

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 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), $E(Y_t^T - Y_t^C | D_t = 1)$. With $t = 1$ denoting SY 2004/05 and $t = 0$ denoting SY 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 ($B_1 = B_0$), which is also called the parallel trends assumption, 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

[16] In SY 2002/03, total grade 5 enrollment was 94.1 percent of the total grade 4 enrollment in TEEP schools on average, compared with 95.4 percent in non-TEEP schools; and the total grade 6 enrollment was 94.6 percent of the total grade 5 enrollment in TEEP schools on average, compared with 95.5 percent in non-TEEP schools.

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; that is, $[(Y_1^C - Y_0^C) \perp D|X]$. 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

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

where N_1 is the number of treated schools and W_{ij} is the weight corresponding to school *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 kernel method instead of the more commonly used nearest-neighbor matching to obtain valid bootstrapped standard errors (Abadie and Imbens 2006, 2008).

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

Since ATET can be estimated consistently only 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.

5. Average Treatment Effects

In the estimation, we merged NAT grade 4 in SY 2002/03 and NAT grade 6 in SY 2004/05 using elementary schools in SY 2002/03.[17] Although the selection of TEEP is based on province-level poverty indicators summarized in the Social Reform Agenda, we conjecture that income distributions overlap between TEEP and non-TEEP school districts (see Figures 3.1 and 3.2). In our matching estimation, we control for the interactions of municipality income category and regional dummies, as well as school-level initial conditions including pupil–teacher ratio, grade 4 total enrollment, number of multi-grade classes, and proportion of locally funded teachers. In the Philippine context, local income level not only summarizes broad socioeconomic 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 (see, for example, Yamauchi 2005). It also controls local labor market conditions.

The first-stage logit regression result is reported in Table 5.1. The dependent variable is 1 if the school is located in a TEEP area and zero otherwise. The results show that income categories, distinguished by regions, significantly explain TEEP placement. Except for income category 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 category 5 is negative. In other regions, western and eastern Visayas, the income effect is monotonic throughout all income classes.

[17] Our analysis pertains only to elementary schools in SY 2002/03, which offered grades 1 to 6. To maintain a valid cohort, we dropped primary schools, where only grades 1 to 4 are taught.

The pseudo R-squared of the logit regression is 0.22, which suggests plausible explanatory power. The PS of each observation is estimated based on the regression. Figure A.1, in the Appendix, plots densities of the estimated PS in the 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 A.1 displays simple differences of the explanatory variables between the treatment and control groups in the untrimmed sample and the PS weighted and trimmed samples. Although simple differences between the groups are large and statistically significant in the untrimmed sample, trimming and matching based on the propensity score eliminates all significant differences.

In Table 5.2, we report the estimation results on ATET of TEEP. We examine changes in overall and mathematics NAT scores from grade 4 in SY 2002/03 to grade 6 in SY 2004/05.[18] 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 and PS matching (weighted regression) and DD and PS matching (kernel), respectively. The two methods give close results, which suggests that TEEP has significant impacts on both overall and mathematics scores. The magnitude is about 4 overall and 5 for mathematics.[19] In other words, TEEP attributes to an increase of

[18] Mathematics is the only common subject that was tested by all schools in the two grades. Overall score is the summation of scores of all the subjects being tested.
[19] Interestingly, mathematics scores increased more than the overall scores, though the initial level of the average mathematics score was higher than the overall score.

about 6 percent in the overall test score and 8 percent in the mathematics score on average.[20] The impact is not trivial over the two-year period. If the marginal impact is constant across all grades (thus if extrapolated), the total effect of TEEP over six years (if students are exposed to TEEP in the entire elementary school period) would be a score increase of about 12 to 15 points. This magnitude of performance improvement is substantial. We note that the DD and PS matching 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) would tend to have a lower trend in NAT than non-TEEP schools if TEEP were not in place.

To check robustness of our results, we use alternative matching methods based on the nearest neighbor and 5 nearest neighbors and matching based on radius less than 0.05 and radius less than 0.1. For each of the matching methods, we use two trimming methods. One method is the same as described in section 4. The other method trims off treatment observations whose PS score is higher than the maximum or less than the minimum the PS score of the controls. The estimation results are summarized in Table A.2. The results suggest that TEEP effects are statistically significant for both overall score and math score for each combination of matching and trimming methods. The estimated magnitudes are close to those from our main result. The ATET ranges from 3.60 to 4.78 for overall score and from 4.82 to 7.07 for math score over these combinations.

[20] This is computed by dividing the estimated ATET of TEEP by the counterfactual average score of the trimmed treatment group in SY 2004/05.

6. Component-wise Analysis

The previous analysis suggests that TEEP, as a whole, has a significant effect on school performance. Because TEEP is a combination of several components, in this section we explore how each component contributes to school performance.[21] To do so, we specify the empirical model as

$$\Delta H = \alpha + \beta_1 \Delta Textbook + \beta_2 \Delta Training + \beta_3 \Delta Building + z\gamma + \varepsilon \quad ,$$

where ΔH is the change in human capital (measured by test scores) from SY 2002/03 to SY 2004/05 . $\Delta Textbook$, $\Delta Training$, and $\Delta Building$ are TEEP investments in textbooks, teacher training, and building, respectively, that are expected to benefit the cohort under study.[22] Investments in textbooks include those for grades 4, 5, and 6 separately. Investments in training include instruction training and subjective training of teacher. Investments in building refer to the number of new school constructions and new renovations. z is a vector of the initial district- and school-level conditions including the interactions of municipality-level income categories and regional dummies, pupil–teacher ratio, grade 4 enrollment, number of multi-grade classes, and proportion of local funded teachers. We note that the initial human capital and TEEP investments are potentially

[21] School building construction and renovation projects were grouped into two types: local government led and school principal led. In school principal-led projects, schools can propose construction/renovation projects under the condition that they contribute to the funding and actual implementation (e.g., fund raising with communities, labor services).

[22] For example, grade 4 textbook refers to the textbooks distributed to grade 4 in SY 2002/03. The grade 4 textbook distributed to grade 4 in SY 2003/04 is not counted because it did not benefit our cohort. For each grade, the value of new textbooks distributed was divided by the number of students enrolled in the corresponding grade. The total enrollment changes over grades since some students cannot progress to the next grade (i.e., repeating or dropout). Some students also transfer from primary to elementary schools at grade 5. We decided to use the enrollment of each grade to normalize the value of new textbooks, but we could also divide it by the initial grade-4 enrollment. Preliminary analysis showed that the results are qualitatively the same.

complementary (and thus not separable), but we assume that the initial school conditions are sufficient to control such heterogeneities in the intervention effect.

The results are presented in Table 6.1, both for the entire sample and for the TEEP-only sample. 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 also significant, although it has a much lower magnitude. Third, instructional training seems to have a greater positive effect on student performance than subject-wise training (mathematics, English, and so forth). The latter has a negative effect on student performance, at least in the short run, probably because teachers have to use their teaching time to receive training.

Table A.3 in appendix shows the determinants of the above TEEP investments. Each TEEP investment was regressed on region dummies, income class categories (and their interactions), pupil–teacher ratio, grade 4 enrollment, number of multi-grade classes, and proportion of local funded teachers. First, Central and Western Visayas generally received smaller amounts of the TEEP textbooks than Eastern Visayas. Second, except new building constructions, large pupil teacher ratios are negatively associated with TEEP investments in per-student terms. Third, though we do not find general patterns in the effects of municipality-level income classes. Therefore, the above results do not generally support the proposition that the allocation of TEEP investments was pro-poor.

The component-wise impact analysis has 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 the impacts of the textbooks are already reflected in their NAT scores at SY 2002/03 (grade 4). Second, though we have information on school building project contract values, we use the number of new constructions and renovations because the contract value aggregates both types and we also conjecture that the impacts are different between new constructions and renovations. These conjectures were supported in preliminary analyses.

Finally, in this study, we did not explicitly assess school-based management, mainly because we did not find appropriate input measures and variations. The batch plan was not strictly implemented especially in the first and second batch groups (that is, they were mixed in reality, depending on the updated preparedness at the division level). This soft component is thought to improve the overall effectiveness of physical investments and teacher training.

7. Conclusion

This paper provided evidence from the Philippines that both physical and soft components of public school education investments significantly increased national achievement test scores. If the marginal effect is constant across all grades, our estimates indicate that the six-year exposure increases the score by about 12–15 points. Our study also showed that the performance in mathematics is more positively responsive to education reform and investments than other subjects. Our results, however, also depend on the validity of various assumptions made in our methodological approach.

Second, we also found evidence that early-stage investments improve student performance at later stages in the elementary school cycle. The distribution of grade 4 textbooks is shown to increase subsequent student test scores more than grade 5 or grade 6 textbooks do. This is not surprising due to the cumulative nature of knowledge acquisition (not just in education), but this dynamic production cannot be identified without exogenous variations in the inputs. Our results imply that improved educational quality at the elementary school stage has positive impacts on educational progress at later stages.

The above findings, when combined with evidence in the literature, imply that public investments in elementary education likely have positive longer term impacts on education performance at the subsequent stages: for example, progression to high schools and colleges and academic performance. If so, social returns to an early-stage investment can be greater than what the current study seems to show. This argument justifies large public investments to improve school quality at the early stage of public education, because the cumulative benefits are gradually realized at later stages in the education system and labor markets.

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 an ability-screening hypothesis that private schools screen high-ability students but their actual schooling investments are not contributing to productivity increase (see, for example, 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, A., and G. W. Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* 74: 235–267.

Alberto Abadie & Guido W. Imbens, 2008. "On the Failure of the Bootstrap for Matching Estimators." *Econometrica* 76(6): 1537-1557.

Alderman, H., J. R. Behrman, V. Lavy, and R. Menon. 2001. "Child Health and School Enrollment: A Longitudinal Analysis." *Journal of Human Resources* 36: 185–205.

Alderman, H., J. Hoddinott, and B. Kinsey. 2006. "Long Term Consequences of Early Childhood Malnutrition." *Oxford Economic Papers* 58: 450–474

Behrman, J. R., and A. B. Deolalikar. 1991. "School Repetition, Dropouts, and the Rates of Return to Schooling: The Case of Indonesia." *Oxford Bulletin of Economics and Statistics* 53: 467–480.

Chen, S., R. Mu, and M. Ravallion. 2009. "Are There Lasting Impacts of Aid to Poor Areas?" *Journal of Public Economics* 93: 512–528.

Crump, R. K., V. J. Hotz, G. W. Imbens, and O. A. Mitnik. 2009. "Dealing with Limited Overlap in Estimation of Average Treatment effects." *Biometrika* 96: 187–199.

Cunha, F., J. J. Heckman, L. Lochner, and D. V. Masterov. 2006. "Interpreting the Evidence on Life Cycle Skill Formation." In *Handbook of the Economics of Education,* Volume 1, edited by E. A. Hanushek and F. Welch, 697–812, Elsevier, North-Holland, Amsterdam, Holland.

Hirano, K. H., G. W. Imbens, and G. Ridder. 2003. "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score." *Econometrica* 71: 1161–1189.

Khattri, N., C. Ling, and S. 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. Washington D.C.

Maluccio, J. A., J. Hoddinott, J. R. Behrman, R. Martorell, A. R. Quisumbing, and A. D. Stein. 2009. "The Impact of Improving Nutrition during Early Childhood on Education among Guatemalan Adults." *Economic Journal* 119: 734–763.

Ramos, F. V. 1995. The Social Reform Agenda: Winning the Future. Quezon City, Philippines: Republic of the Philippines.

Republic of the Philippines, Congress. 1998. Republic Act 8425 (Social Reform and Poverty Alleviation Act).

Republic of the Philippines, Department of Finance. 2001. Department Order No. 32-01 (effective November 20, 2001).

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.

World Bank, 2007, *Project Completion Report: Third Elementary Education Project*, Washington D.C.: World Bank.

Yamauchi, F. 2005. "Why Do Schooling Returns Differ? Screening, Private Schools, and Labor Markets in the Philippines and Thailand." *Economic Development and Cultural Change* 53: 959–981.

Yamauchi, F. 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 3.1—Summary of NAT test scores for TEEP and non-TEEP schools, SY 2002/03 and SY 2004/05

	TEEP		Non-TEEP			
	Mean	s.d.	Mean	s.d.	Diff.	sig.
Overall score in 2002/03	46.97	14.67	44.41	13.50	2.56	***
Math score in 2002/03	48.39	17.96	45.82	16.73	2.57	***
Overall score in 2004/05	63.71	13.43	59.80	12.86	3.91	***
Math score in 2004/05	66.03	16.62	62.23	16.68	3.80	***

Source: National Achievement Test database, various years.

Note: s.d. = standard deviation.

Table 3.2—Percentage of TEEP schools in the Visayas region by the initial implementation timing

	SY 2000/01	SY 2001/02	SY 2002/03	SY 2003/04	SY 2004/05	SY 2005/06
New construction and renovation projects	6%	22%	49%	63%	84%	86%
Grade 1 textbook distribution	76%	76%	81%	100%	100%	100%
Grade 2 textbook distribution	76%	76%	81%	100%	100%	100%
Grade 3 textbook distribution	76%	76%	81%	81%	81%	100%
Grade 4 textbook distribution	76%	76%	81%	100%	100%	100%
Grade 5 textbook distribution	76%	76%	81%	100%	100%	100%
Grade 6 textbook distribution	69%	69%	74%	100%	100%	100%
Training program of teachers	31%	99%	100%	100%	100%	100%

Source: TEEP investment database (the authors' survey), and Division Education Development Plan database

Table 5.1—Logit estimation of TEEP placement

TEEP	Coeff.	
Central Visayas	-2.163***	(0.211)
Western Visayas	-2.518***	(0.226)
Income 2	1.168***	(0.310)
Income 3	1.872***	(0.367)
Income 4	0.306	(0.190)
Income 5	0.142	(0.186)
Central Visayas × Income 2	-1.163***	(0.421)
Central Visayas × Income 3	-1.267***	(0.423)
Central Visayas × Income 4	0.332	(0.259)
Central Visayas × Income 5	-1.977***	(0.388)
Western Visayas × Income 2	-0.610	(0.398)
Western Visayas × Income 3	-1.081**	(0.424)
Western Visayas × Income 4	1.279***	(0.263)
Western Visayas × Income 5	0.954***	(0.312)
Pupil–teacher ratio (both local and national)	-0.00818*	(0.00434)
Grade 4 total enrollment (in ages 6 to 11)	-0.00766***	(0.00141)
Number of multigrade classes	-0.0412	(0.0402)
Proportion of local funded teachers	0.233	(0.595)
Constant	1.294***	(0.212)
Number of observations	4222	
Pseudo R2	0.219	

Note: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Standard errors in parentheses

Table 5.2—Impacts of TEEP on school performance

	Untrimmed sample, simple DD				
	Treated diff	Control diff	**DD**	se	sig.
Overall score	16.74	15.39	**1.35**	0.87	
Math score	17.64	16.41	**1.24**	1.09	
Number of observations	1774	2448			
	Trimmed sample, DD+PS weighted regression				
	Treated diff	Control diff	**DD**	se	sig.
Overall score	16.05	12.21	**3.84**	1.13	***
Math score	16.94	11.79	**5.15**	1.47	***
Number of observations	1541	2422			
	Trimmed sample, DD+PS weighted kernel				
	Treated diff	Control diff	**DD**	se	sig.
Overall score	16.05	12.33	**3.72**	1.04	***
Math score	16.94	12.04	**4.90**	1.29	***
Number of observations	1541	2422			

Note: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

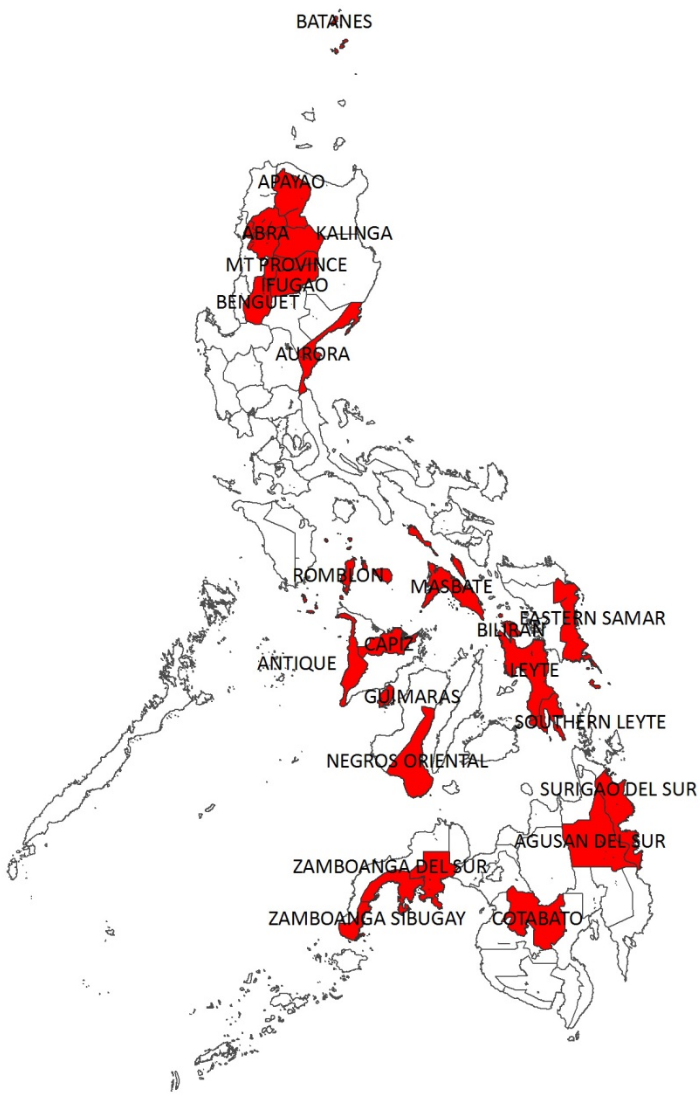
Table 6.1—Estimation results of component analysis, dependent variables being change in mathematics score and overall score

	All Sample				TEEP only			
	Overall Score		Mathematics Score		Overall Score		Mathematics Score	
Grade 4 textbooks (peso/pupil)	0.0418***	(0.00697)	0.0142**	(0.00636)	0.0333***	(0.00542)	0.0129**	(0.00503)
Grade 5 textbooks (peso/pupil)	-0.00726	(0.00474)	-0.000362	(0.00497)	-0.00575	(0.00386)	-0.00112	(0.00395)
Grade 6 textbooks (peso/pupil)	-0.00289	(0.00500)	-0.00243	(0.00494)	-0.00340	(0.00385)	-0.00251	(0.00373)
Instructional training (man-hours/pupil)	0.487**	(0.229)	0.327*	(0.190)	0.427**	(0.177)	0.267*	(0.156)
Subject training (man-hours/pupil)	-0.849***	(0.327)	-0.590*	(0.302)	-0.619**	(0.260)	-0.406	(0.251)
New constructions (number in SY 2003/04)	5.756***	(1.924)	5.316***	(1.970)	5.390***	(1.111)	5.010***	(1.116)
New renovations (number in SY 2003/04)	1.490***	(0.471)	1.199**	(0.489)	1.116***	(0.330)	0.884**	(0.372)
Central Visayas	7.111**	(3.265)	-3.695	(3.994)	3.154	(2.720)	-3.224	(3.912)
Western Visayas	-0.613	(3.398)	-19.41***	(3.356)	-0.254	(2.784)	-14.18***	(2.878)
Income 2	4.176	(4.016)	2.893	(4.557)	4.082	(3.445)	2.474	(3.779)
Income 3	-1.293	(3.319)	-2.525	(3.284)	-0.591	(2.716)	-1.398	(2.753)
Income 4	-0.654	(3.298)	-0.946	(3.507)	-1.027	(2.676)	-1.509	(2.967)
Income 5	2.168	(2.967)	1.181	(3.151)	1.441	(2.447)	0.780	(2.697)
Central Visayas × Income 2	-1.530	(4.638)	-2.906	(5.654)	-0.736	(4.044)	-4.757	(5.355)
Central Visayas × Income 3	-1.758	(4.510)	-2.143	(4.792)	-1.156	(3.676)	-1.751	(4.235)
Central Visayas × Income 4	0.394	(4.020)	-4.268	(5.453)	0.703	(3.245)	-3.634	(4.853)
Central Visayas × Income 5	0.0249	(3.697)	-0.552	(5.529)	0.328	(3.130)	-1.276	(4.406)
Western Visayas × Income 2	-0.623	(4.983)	8.271*	(4.311)	0.0490	(4.229)	6.239	(3.952)
Western Visayas × Income 3	1.083	(4.122)	16.73***	(4.913)	0.597	(3.312)	11.69***	(3.978)
Western Visayas × Income 4	1.006	(3.981)	13.64***	(4.425)	2.310	(3.233)	11.89***	(3.654)
Western Visayas × Income 5	2.199	(4.481)	10.76***	(4.088)	2.551	(3.526)	9.895***	(3.358)
Pupil teacher ratio	-0.118**	(0.0493)	-0.128*	(0.0760)	-0.0990**	(0.0403)	-0.156**	(0.0625)
Grade 4 total enrollment	0.0473***	(0.00991)	0.0576***	(0.0184)	0.0464***	(0.00804)	0.0613***	(0.0149)
Number of multi-grade classes	-0.456	(0.373)	-0.117	(0.605)	-0.504*	(0.283)	0.160	(0.462)

Proportion of local funded teachers	-11.90*	(6.780)	-6.336	(14.29)	-8.641	(5.531)	-9.587	(11.86)
Constant	15.52***	(3.289)	21.67***	(3.701)	15.25***	(2.659)	20.98***	(3.060)
Number of observations	3905		1471		3905		1471	
R-squared	0.059		0.088		0.060		0.113	

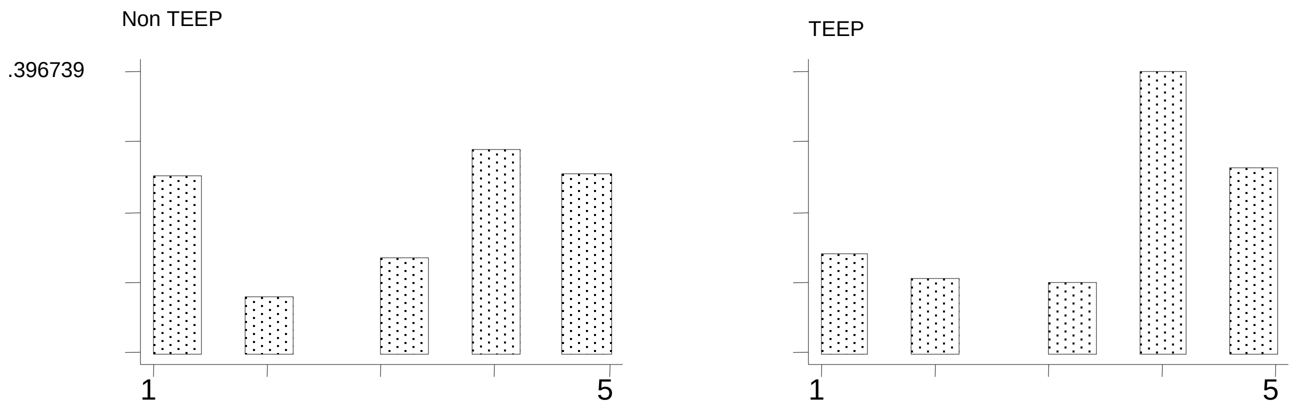
Pesos are in Philippine pesos, PHP. Standard errors are in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level

Figure 2.1—Map of TEEP and non-TEEP divisions in Philippines (TEEP areas are in red)



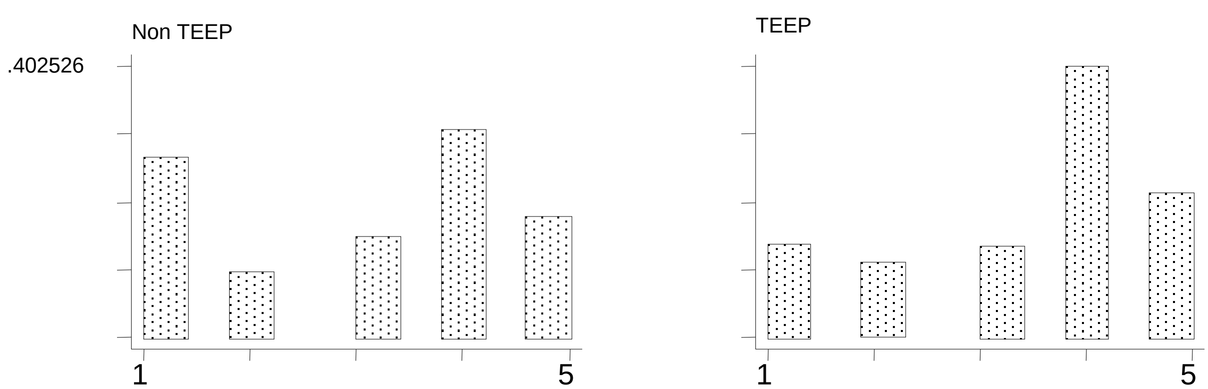
Source: The authors' calculation

Figure 3.1—Histogram of school districts by income category for TEEP and non-TEEP groups



Source: Census 2000 Municipality Income Classifications

Figure 3.2—Histogram of sampled schools by income category for TEEP and non-TEEP groups



Source: Census 2000 Municipality Income Classifications

Appendix

Table A.1 Balance check

variables	Untrimmed sample, simple DD			Trimmed sample, PS weighted regression			Trimmed sample, PS weighted kernel		
	Diff.	s.e.	Sig.	diff.	s.e.	sig.	diff3	se3	sig3
Central Visayas	-0.284	0.047	***	-0.003	0.046		-0.012	0.044	
Western Visayas	-0.147	0.050	***	0.000	0.055		0.000	0.058	
Income 2	0.002	0.031		0.002	0.017		-0.003	0.021	
Income 3	0.000	0.040		0.000	0.035		-0.004	0.031	
Income 4	0.104	0.050	**	0.004	0.062		0.024	0.064	
Income 5	0.022	0.039		-0.001	0.054		-0.001	0.049	
Central Visayas × Income 2	-0.024	0.015		0.000	0.010		-0.002	0.009	
Central Visayas × Income 3	-0.026	0.026		-0.001	0.025		-0.002	0.025	
Central Visayas × Income 4	-0.047	0.032		-0.002	0.032		-0.001	0.029	
Central Visayas × Income 5	-0.101	0.020	***	0.000	0.005		-0.002	0.006	
Western Visayas × Income 2	-0.032	0.019		0.000	0.014		-0.004	0.015	
Western Visayas × Income 3	-0.040	0.027		0.000	0.025		-0.004	0.025	
Western Visayas × Income 4	0.021	0.039		0.000	0.047		0.007	0.044	
Western Visayas × Income 5	-0.008	0.014		-0.001	0.014		0.002	0.014	
Pupil–teacher ratio	-2.215	0.758	***	-1.075	0.847		-1.282	0.841	
Grade 4 total enrollment	-7.381	1.323	***	0.716	1.194		0.584	1.098	
Number of multi-grade classes	0.134	0.049	***	-0.039	0.076		-0.042	0.083	
Proportion of local funded teachers	-0.005	0.003		-0.001	0.004		0.000	0.004	
Number of observations	4222			3963			3963		

Note: DD: Double difference, PS: Propensity score, se: Standard errors, diff: mean-difference, *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Table A.2 Impacts of TEEP on school performance: robustness check using other matching methods

	Matching based on the nearest neighbor					
	Trimming method 1			Trimming method 2		
Overall score	4.30	(1.10)	***	3.84	(1.21)	***
Math score	7.07	(1.62)	***	6.29	(1.70)	***
	Matching based on 5 nearest neighbors					
	Trimming method 1			Trimming method 2		
Overall score	4.78	(1.21)	***	3.94	(1.17)	***
Math score	6.72	(1.55)	***	5.59	(1.54)	***
	Matching based on radius <0.05					
	Trimming method 1			Trimming method 2		
Overall score	4.64	(1.14)	***	3.88	(1.12)	***
Math score	6.25	(1.47)	***	5.12	(1.47)	***
	Matching based on radius<0.1					
	Trimming method 1			Trimming method 2		
Overall score	4.29	(1.11)	***	3.60	(1.12)	***
Math score	5.95	(1.41)	***	4.82	(1.45)	***

Trimming method 1: Trimming off treatment observations whose PS score is higher than the maximum or less than the minimum the PS score of the controls.

Trimming method 2: as described in section 4.

Standard errors (in the parentheses) are not bootstrapped.

Table A.3 Determinants of TEEP Investments

	Gr4	Gr5	Gr6	Instr.Training	Sub.Training	New const.	Renovation
Central Visayas	-122.6***	-114.5***	-113.6***	-0.581	-0.800	0.00253	0.0414***
Western Visayas	-131.6***	-149.6***	-155.2***	-1.943***	-1.388***	0.00269	-0.000485
Income 2	95.11**	128.9**	113.0*	2.230	1.572	-0.00169	-0.0237
Income 3	103.3***	147.3***	108.9***	-0.208	-0.371	0.0000698	-0.00262
Income 4	8.378	19.61	12.78	-0.224	0.184	-0.0000723	0.00456
Income 5	18.83	41.20	25.45	0.486	-0.249	-0.00117	-0.0216**
Central Visayas × Income 2	-99.87**	-131.6*	-88.37	-3.212	-1.724	0.000852	0.0113
Central Visayas × Income 3	-104.9***	-123.2**	-90.98*	-0.260	0.336	-0.000743	-0.00351
Central Visayas × Income 4	-12.99	14.10	15.09	-0.0821	0.583	-0.00173	-0.0205*
Central Visayas × Income 5	-32.74	-85.03**	-70.03*	-1.806*	-0.431	-0.00235	-0.0349**
Western Visayas × Income 2	-66.03	-74.30	-62.21	-1.605	-1.209	0.0215	0.182
Western Visayas × Income 3	-106.4***	-87.34*	-41.98	1.252	0.993*	0.0662	0.148
Western Visayas × Income 4	2.281	78.86*	84.32**	1.241**	0.517	0.0439**	0.477***
Western Visayas × Income 5	9.294	83.14	94.75*	0.860	1.275*	0.231**	0.598**
Pupil teacher ratio	-1.090***	-1.523**	-1.102*	-0.0348***	-0.0273***	-0.000163	-0.00527**
Grade 4 total enrollment	-0.0486	-0.209***	-0.269***	-0.00252	-0.00219*	-0.0000463	0.000231
Number of multi-grade classes	-1.692	4.940	4.868	0.0870	0.0134	0.00349	0.0362***
Proportion of local funded teachers	-41.44**	47.85	48.80	0.652	0.388	-0.0153	0.375
Constant	171.0***	215.2***	204.9***	3.185***	2.451***	0.00503	0.136**
Number of obeservatkions	3905	3905	3905	3905	3905	3905	3905
R-squared	0.469	0.241	0.215	0.114	0.081	0.053	0.069

*** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level

Figure A.1—Plot of estimated propensity scores for schools in non-TEEP and TEEP areas