Analysis Plan for Early Childhood Teacher Development using an On-site Training Approach in Rural Thailand

September 30, 2022

The analysis plan for this study is as follows.

1 How to Construct Outcomes

Primary outcomes of this study are derived from the following three measures:

1. Child development measuring by DENVER II.
2. Picture Memory test, taken from the Wechsler Preschool & Primary Scale of Intelligence (WPPSI).
3. Social-emotional skills based on Strengths and Difficulty Questionnaire (SDQ).

The first two measures will be directly assessed at both baseline and endline surveys while the parent will be asked Strengths and Difficulty Questionnaire (SDQ) in the endline only.

1.1 DENVER II and Picture Memory

We will construct five measures of child development based on DENVER II, namely gross motor ($j = 1$), language ($j = 2$), fine motor-adaptive ($j = 3$), personal-social ($j = 4$), and total ($j = 5$). We will construct a single measure of child development based on picture memory test, namely working memory ($j = 6$).

We will construct standardized scores of child development in two different ways.

1. Standardized Scores using Raw Scores: We will first generate a raw score for each of the skills by counting the number of items the child can pass/correctly answer. Note that the total score ($j = 5$) is simply the sum of all the first four scores. Let Y_{ij} denote the raw score of skill j for child i. We will then derive age-standardized scores for each skill j using kernel-weighted local polynomial smoothing up to the third degree polynomial (as in, Attanasio et al., 2020). We will first perform this age-standardization for the baseline, and then use the predicted mean and standard deviation for the baseline to perform age-standardization for the end line. That is, the standardized scores for this case will be relative to the mean and SD of the baseline.

2. Standardized Scores using Item Response Theory (IRT): The only difference from the first method is that we will use latent scores from IRT instead of the raw scores. This approach is similar to the one used in Zhou et al. (2022). We will first derive adjusted raw score by performing IRT analysis on each skills. This method will give latent scores for every skill sets. We will then derive age-standardized scores for each skill j using kernel-weighted local polynomial smoothing up to the third degree polynomial.

1.2 SDQ

Strengths and Difficulty Questionnaire (SDQ) contains 5 sub-scales, namely conduct problems, emotional symptoms, hyperactivity/inattention, prosocial behaviour and peer-relationship problems. We will generate raw scores for each sub-scale by setting the score for each item in such a way that a higher score means less problematic behavior. That is, for negative questions, we will set the score for each item as follows: Not true = 2, Somewhat true = 1, True = 0, while the opposite will be applied for positive questions. We will then derive age-standardized scores for each skill j using kernel-weighted local polynomial smoothing up to the third degree polynomial. The next step is to estimate the latent factor, called social-emotional skill, using a confirmatory factor analysis (CFA) approach (e.g., Gorsuch, 1983; Thompson, 2004).

1.3 Teaching Quality

The research team will send early childhood education experts to visit all the centers. They will also collect information regarding teaching quality using a classroom observation record, created by the research team. We will apply exploratory factor analysis (EFA) to determine the number of latent factors relevant for teaching quality and to group items together. We will estimate the latent factors using a confirmatory factor analysis (CFA) approach

2 Econometric Specifications

2.1 Impacts of the Program on Child Skills

Main outcome variables in this study are in the form of daily learning gains, measuring by the value-added of test scores divided by school days between the baseline and endline.

More formally, let θ^0_{ij} and θ^1_{ij} denote standardized scores for skill j of student i at the baseline and the endline, respectively. The daily learning gain for skill j of student i is defined by $\frac{\theta^1_{ij}-\theta^0_{ij}}{\tau_i}$, where τ_i is the number of school days between the baseline and endline tests (excluding the days that treated teachers were participating in the on-site training for the treatment group). This learning gain reflects the effectiveness of the teacher/classroom with respect to child skills. Defining the daily learning gain this way also helps mitigate a potential bias arising from the difference of school days between the treatment and control groups, (as discussed in Loyalka et al., 2019). On the other hand, this approach limit our analysis to skills that were measured both at the baseline and endline tests.

The benchmark model estimates the following linear model:

$$\frac{\theta^1_{ij}-\theta^0_{ij}}{\tau_i} = \alpha + \beta T_i + \boldsymbol{\gamma}\boldsymbol{X}_i + \varepsilon_{ij}, \tag{1}$$

where T_i is a dummy variable indicating whether student i attended a treatment school during the experiment; $\boldsymbol{X}_i$ is a vector of control variables; and ε_i is an error term. The key parameter of interest, β, estimates the intent-to-treat effect (ITT).

To deal with the non-compliance problem, we estimate the following treatment-on-the-treated effect (TOT) using the treatment dummy T_i as the instrument.

$$\frac{\theta_{i1}^s - \theta_{i0}^s}{\tau_i^{01}} = \alpha + \beta C_i + \boldsymbol{\gamma} \boldsymbol{X}_i + \varepsilon_{ij}, \tag{2}$$

where C_i is a dummy variable indicating that child i's teacher(s) received the on-site training or not, and is instrumented by the treatment dummy T_i.

For the social-emotional skill, we cannot use the daily learning gain as the outcome since it will be measured at the endline only. We, therefore, use the following specifications for ITT and TOT of social-emotional skill, respectively:

$$\theta_{ij}^1 = \alpha + \beta T_i + \boldsymbol{\gamma} \boldsymbol{X}_i + \varepsilon_{ij}, \tag{3}$$

$$\theta_{ij}^1 = \alpha + \beta C_i + \boldsymbol{\gamma} \boldsymbol{X}_i + \varepsilon_{ij}, \tag{4}$$

where the second specification will be instrumented by the treatment dummy T_i.

2.2 Impacts of the Program on Teaching Quality

We will estimate the impacts of the program on teacher quality indices. This analysis will shed light on the mechanism(s) through which the program may affect child outcomes.

Let Q_{kl} denote teaching quality index k for classroom l, T_l denote a dummy variable indicating whether classroom l is in a treatment center or not, and C_l denote a dummy variable indicating whether a teacher in classroom l received the on-site training or not. We will estimate both ITT and TOT effects for teacher quality indices using the following specifications.

$$Q_{kl} = \alpha + \beta T_l + \boldsymbol{\gamma} \boldsymbol{X}_l + \varepsilon_{kl}, \tag{5}$$

$$Q_{kl} = \alpha + \beta C_l + \boldsymbol{\gamma} \boldsymbol{X}_l + \varepsilon_{kl}, \tag{6}$$

where the second specification will be instrumented by the treatment dummy T_l and $\boldsymbol{X}_i$ is a vector of control variables related to the classroom. We will also control for evaluator fixed-effect (in $\boldsymbol{X}_l$).

2.3 Heterogeneous Effects

We will estimate heterogeneous effects using the standard regression model with interaction terms. The heterogeneous variables of interest include child gender, household wealth, parental education, teacher education etc.

References

Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., and Rubio-Codina, M. (2020). Estimating the production function for human capital: results from a randomized controlled trial in Colombia. *American Economic Review*, 110(1):48–85.

Gorsuch, R. L. (1983). *Factor Analysis*. Psychology Press, New York.

Loyalka, P., Popova, A., Li, G., and Shi, Z. (2019). Does teacher training actually work? evidence from a large-scale randomized evaluation of a national teacher training program. *American Economic Journal: Applied Economics*, 11(3):128–54.

Thompson, B. (2004). *Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications*. Applied Psychological Measurement, Washington, DC.

Zhou, J., Heckman, J. J., Liu, B., and Lu, M. (2022). The impacts of a prototypical home visiting program on child skills.