**Impact of early childhood school intervention on enrollment and learning outcomes: Evaluation of a public program in India**

Nilanjan Bhattacharya[♦], Debayan Pakrashi[♥], Sarani Saha[♠] and Soham Sahoo[ℵ]


## Abstract

We evaluate the impact of introducing a pre-primary schooling program in government schools in the Indian state of West Bengal in 2013 on children's early enrolment in schools and subsequent test scores. Using double difference, triple difference, and synthetic control methodologies, we find that the program significantly increased enrolment in the pre-primary sections of the government schools. However, the rise in enrolment did not translate into improved performance of the students. Analyzing the test scores, we find that after the program's introduction, both math and reading scores of treated children did not improve compared to the control group. We attribute this result to the deteriorating physical and learning infrastructure in the state government schools, captured via a decline in the availability of classrooms and teachers.


**Keywords:** pre-primary education; learning outcomes, enrolment

**JEL Classification:** I21, I25, I28


♦ Department of Economic Sciences, Indian Institute of Technology Kanpur. Email: nilanjanb1@gmail.com

♥ Department of Economic Sciences, Indian Institute of Technology Kanpur and Economic Research Unit, Indian Statistical Institute Kolkata. Email: pakrashide@gmail.com

♠ Department of Economic Sciences, Indian Institute of Technology Kanpur, India. Email: sarani@iitk.ac.in

ℵ Centre for Public Policy, Indian Institute of Management Bangalore, India. Email: soham.sahoo@iimb.ac.in

## 1. Introduction

Good quality early childhood education (ECE) is important in ensuring fair education opportunities for everyone. ECE can lead to better learning outcomes (Berlinski, Galiani, and Gertler 2009; Britto et al. 2017), equitable learning opportunities for marginalized groups (Berlinski, Galiani, and Manacorda 2008; Heckman 2013; Elango et al. 2015), improved health outcomes (Elango et al. 2015), and even long-term economic benefits extending to adulthood (Heckman et al. 2010; Gertler et al. 2014; García et al. 2017). These findings have given rise to the concept of dynamic complementarity in skill formation. Children exposed to early childhood programs may benefit more from later human capital investments (Cunha and Heckman 2007). Given the benefits, it is obvious that there has been a rapid increase in ECE programs globally in the recent past (Nores and Barnett 2010; Behrman and Urzúa 2013; Cascio 2015; Sayre et al. 2015; Wotipka et al. 2017). Providing decent early childhood care, and pre-primary education by 2030, as envisioned in target 4.2 of the Sustainable Development Goal 4 (UN 2015), also echoes its importance.

Despite the high return to early childhood investment in education, an estimated 175 million, which corresponds to almost fifty percent of the world's children between 3 and 6 years (pre-primary age group), are deprived of pre-primary schooling (UNICEF 2019). Only one in every five children has been exposed to pre-primary education in developing and low-income countries (UNICEF 2019). This deprivation is an obstacle to achieving the children's developmental potential. The analysis of an ECE intervention in a low and middle-income context is essential as these countries account for most of the children in the world.

In this paper, we evaluate the causal impact of a government-run free pre-primary program implemented in 2013 on children's enrolment and test scores in the Indian state of West Bengal. The effect of such a program is more relevant in the context of India, where quality ECE programs are not available to the young masses, particularly children coming from underprivileged families. To explore the pathway, we also analyze if schools' infrastructure has any role that ultimately affects the children's learning outcomes. Our study also aligns with India's recently announced, 'National Education Policy 2020' (NEP 2020). In India's existing 10+2 education structure, children in the age group of 3-6 are not included. Presently, a child is admitted to class 1 at the age of 6. The NEP 2020 proposes a 5+3+3+4 framework, covering children from age three. It sets a strong foundation for Early Childhood Education (ECE) aimed at the holistic development of the child. Focussing on a government-run free pre-primary program, we analyze the effectiveness of universal provisioning of quality early childhood development as envisioned in the NEP 2020.

In particular, we analyze the causal impact of a government-run, pre-primary program on enrolment and learning outcomes by exploiting the exogenous variation in exposure to pre-primary schooling brought about by implementing the program in 2013. We employ double-difference and triple-difference methodologies to estimate the effects of the program on enrolment and test scores. We have also used the synthetic control methodology to rule out any bias in results that can occur due to the ad-hoc nature of the states selected for the triple-difference regressions.

Our main findings suggest that the program has succeeded in increasing enrolment but has no impact on learning outcomes captured via math and reading scores. Results from the difference-in-differences show that the program increased the number of government schools with a dedicated pre-primary section compared to the private schools in West Bengal and government schools in the control states of Bihar, Jharkhand, and Orissa. A similar result for the enrolment of pre-primary students has been found. The change in pre-primary enrolment in the government schools in West Bengal, therefore, showed a massive increase after the program was introduced in 2013 compared to their private counterparts and the control states.

The success story of enrolment, however, did not translate into better learning outcomes. A comparison of pre-primary children between the treated and control cohort shows that treated cohort children from government schools in West Bengal do not achieve any better learning skills when compared to the control cohort.

We attribute this result to the worsening infrastructure of the state government schools in West Bengal. Comparing the number of pre-primary teachers and total teachers between the government schools in West Bengal and the government schools in its neighboring states, we find that number of pre-primary teachers increased. In contrast, the percentage change in the number of total teachers reported a decline in the government schools of West Bengal. The number of pre-primary teachers increased since the existing pool of teachers was assigned to teach at the pre-primary level. Therefore, even though the reported number of pre-primary teachers was high in government schools in West Bengal, the school infrastructure did not complement the higher enrolment in pre-primary by recruiting more teachers, leaving the schools with more students but fewer teachers.

This paper contributes to the growing literature on early childhood education in the context of developing countries in several important ways. First, existing studies have found impactful results by focusing on well-executed ECE programs in developed countries. However, ECE programs in low- and middle-income countries significantly differ from the developed ones (Behrman and Urzúa 2013). In our paper, we find the evidence by focussing on a government-run free pre-primary program in a lower-middle income developing country context. Second, the targeted populations come from disadvantaged backgrounds, which throws the question of the generalizability of the evidence (Baker 2011). In our paper, the program has a free and universal rollout aimed at the general population. Third, existing evidence mostly comes from a few randomized-control trials that face small sample size issues and therefore suffer from the problem of generalizations (Heckman 2011). We eliminate this limitation by using the government-administered DISE (District Information System for Education) and ASER (Annual Status for Education Report) data which are nationally representative databases with large sample sizes. Therefore, the significant contribution of this paper lies in evaluating the government's pre-primary program, which is the first study to causally evaluate a specific early childhood government intervention in the context of India, to the best of our knowledge, and suggest appropriate policy reforms.

The rest of the paper is organized as follows. Section 2 describes the program in detail, followed by a discussion on the datasets and some descriptive statistics in Section 3. The empirical strategy used in the paper is discussed in detail in Section 4. Next, we discuss the results of the paper estimated using the specification mentioned earlier in Section 5. In Section 6, we further analyze the possible reasons for the obtained results and finally conclude the paper with Section 7.

## 2. Description of the Program

India has one of the world's largest education systems, with the number of schools as high as 1.5 million. It also includes 8.5 million teachers, and 250 million students (UNICEF 2018). However, there is a downside to the story when it comes to learning proficiency, as almost half of the primary school children, estimated to be nearly 50 million – could not achieve grade-level standards (NCERT 2017). For example, the percentage of grade 3 children who could read a grade 2 level text was only 27.2% in 2018 (ASER 2018). This learning crisis is mainly common in the initial years of schooling. Most children enter primary school without any prior preparation. In addition, there is widespread inequity in the education system of India due to various socio-economic factors. In rural India, there exists a learning gap between first generation learners with low family income and children from wealthier families with educated parents, by the time they attain the age of seven (Alcott and Rose 2017). A quality early childhood education program can close this learning gap.

Recognizing its importance, section 11 of the 'The Right of Children to Free and Compulsory Education Act, 2009' of India states that "*to prepare children above the age of three years for elementary education and to provide early childhood care and education for all children until they complete the age of six years, the appropriate government may make necessary arrangement for providing free pre-school education for such children*". The national legal framework provided a guarantee through this act, and the government of West Bengal (WB) introduced a free one-year pre-primary education in government schools in the academic session of 2013.[1] According to the new rules, a student aged between 5 and 6 years on the first day of the academic session (i.e., on the 1st of January 2013) would be eligible to take admission in the pre-primary section. Apart from the pre-primary section, the age criteria for admission to different classes were also revised. Before 2013, five-year-old children could get admission in grade 1. However, starting in 2013, only six-year-old students could enroll in grade 1 of the government schools.[2]

The program guidelines also mentioned that separate seating arrangements should be made available to the pre-primary students as far as possible. If, due to the unavailability of space, schools cannot accommodate them in a separate classroom, they could sit with grade 1 students. The existing teachers in government schools should continue teaching until new teachers are recruited. The pre-primary students are also entitled to benefits under the mid-day meal scheme. After having the mid-day meal, pre-primary students are allowed to leave the school.

[1] This policy was implemented in both government owned and government aided schools. We combine these two categories and refer to them as 'government' or 'public' schools throughout the paper. Further details of the policy can be found at: https://wbxpress.com/files/2012/11/Admission_Age.pdf
[2] https://wbxpress.com/files/2012/07/Age_Admission.pdf

## 3. Data & Descriptive Statistics

We use data from the Annual Status of Education Report (ASER), a yearly survey conducted to assess the education status among children in almost all the rural districts of India from 2009 to 2018. The survey covers a random sample of about 20 households from each of the 30 villages selected from the rural districts of India. From each household selected, all children in the age group 5 to 16 are surveyed and administered with the learning test module.

The survey administers basic arithmetic and reading tests and uses the same tools for all children across the states. These test scores have been extensively used in the literature (Chakraborty and Jayaraman 2019; Lahoti and Sahoo 2020; Das and Sarkhel 2023). In the ASER data, the assessment of reading skills has ordinal ranking with an increased level of difficulty—recognition of letters, reading of words, reading a short paragraph (a grade 1 level text), and reading a short story (a grade 2 level text). On the other hand, the arithmetic test is also based on similar levels—recognition of single-digit numbers, recognition of double-digit numbers, subtraction of two-digit numbers with a borrowing, and division of a three-digit number by one digit.

In our analysis, we have considered ASER scores as a measure of learning outcomes.Since these scores are independently collected by ASER, they are standardized and uniform across the schools; hence they have better comparability than any test conducted by the schools themselves. The ASER survey also collects child, household, and village-level information that we use as control variables. Household level characteristics include possession of a cemented house; presence of electricity; possession of a toilet; possession of a television; and the total number of family members. Child-level characteristics include the child's age, and gender. Village-level factors include presence of a private school, a private health clinic, a bank, or a cemented road in the village.

However, ASER does not include any information on whether the child in question is enrolled in the pre-primary section of a school, so it is impossible to identify the children's pre-primary education status. To analyze the program's effect on enrolment at the pre-primary level, we have used District Information System for Education (DISE) database. The DISE dataset gathers detailed information on different school-level characteristics including school infrastructure, facilities, enrolment, and teachers for all the districts. In our analysis, we used DISE data from 2009 to 2017. The sample in our study consists of 541,801 government schools and 75,637 private schools from West Bengal. We also include 1,447,223 government schools and 42,871 private schools from the control states. For learning outcomes, our estimation sample includes 6,550 children in 5-6 years cohort in government schools and 1,755 children of the same cohort in private schools of West Bengal. For the control states, 55,863 children in government schools in the 5-6 years age cohort and 8,108 children in private schools have been considered. The summary statistics are reported in Table 1.

To begin, we use the enrolment and infrastructure data from DISE to see if the program’s implementation in 2013 increased the number of government schools with a pre-primary section and the number of students enrolled in the pre-primary section in West Bengal. See Figure 1.

**[Insert Figure 1]**

Interestingly, we notice an immediate jump in the number of government schools with a pre-primary section and the number of students enrolled in pre-primary for 2013. While the number of government schools with pre-primary drastically increased to almost 60,000, the total number of students enrolled in the pre-primary sections of these government schools recorded a figure of nearly 800,000 in 2013. Next, we compare the number of government and private schools with pre-primary section in West Bengal and pre-primary students enrolled in these schools to the neighboring states of Bihar, Jharkhand, and Orissa to check if these states also reported any similar trends. We consider this group of states as a comparison group to West Bengal since they share their geographical boundary with West Bengal and have many similarities in the social, economic and cultural domain. Figure 1 suggests that the huge expansion of pre-primary education in government schools post-2013 took place only in the state of West Bengal and not in the neighboring states. We did not observe such a massive change in private schools, although there is an increasing trend of pre-primary education in private schools. We also presented child-level learning outcomes using math and reading scores from ASER data in Figure 2. However, we find nothing similar to the pattern in Figure 1.

**[Insert Figure 2]**

Figure 1 further motivates us to dig deeper into estimating the causal impact of the pre-primary program.

## 4. Empirical Methodology

We exploit the exogenous age restriction related to the admission criteria in pre-primary that makes the identification of the causal impact feasible[3]. In particular, we use a difference-in-differences (DD) method to causally estimate the program's impact on pre-primary enrolment and learning outcomes. We use DISE data to assess the impact program on enrolment at the school level and rely on ASER data while focusing on the learning outcomes at the individual level.

The identification of the causal impact of the pre-primary program is challenging due to selection bias. A situation might arise where only motivated parents send their children to pre-

[3] Prior to implementing the pre-primary schooling program, the entry age for government school students was five years for grade I. After the program's introduction, the entry age for children in schools remained the same. However, these five-year-old children were eligible to be enrolled in the pre-primary section of the schools only. They would enjoy a year of pre-primary education before they are promoted to grade I at the age of six.

primary schooling. In such a case, children who attend pre-primary school might possess characteristics that promote better school performance, which could lead to a spurious positive correlation between pre-primary attainment and learning outcomes. However, as the program in discussion over here has a universal roll-out at the state level and the age at entry remains the same, the question of motivated parents sending their children for pre-primary education does not arise. The chance of positive selection bias is, therefore, ruled out. We hypothesize that this one year of extra schooling in the form of pre-primary education for government schools can have an impact on these children's academic achievement, measured using ASER administered reading and math scores.

### *4.1. Difference-in-differences*

The government schools in West Bengal are considered as the treated group as the program was implemented in these schools only. The control group may comprise of the government schools in the neighbouring states that are not exposed to the program. Alternatively, the private schools from West Bengal may also form the control group as the policy was implemented only for the government schools in the state. Accordingly, we have two different specifications for the double difference regression. The specifications are the same for both sets of regressions for enrolment and learning outcomes. The only difference is that the dependent variable for enrolment includes the availability of a pre-primary section and enrolment in pre-primary, which are measured at the school level, whereas the test scores for 5-6 years old individual children are considered as learning outcomes.

In the first specification, the difference between the government schools of West Bengal and the control states after the implementation of the program (i.e., the year 2013) is compared to the same types of schools before the program was implemented. The corresponding difference-in-differences (DD) estimating equation is as follows:

$$(1)\ y_{idst} = \beta_0 + \beta_1.post_t + \beta_2.(post_t \times wb_s) + \vartheta X_{idst} + \partial_{ds} + \gamma_t + (\partial_{ds} \times t) + \epsilon_{idst}$$

Here, $Y_{i\,dst}$ is the outcome variable of interest (availability of pre-primary section in the school and number of students enrolled in pre-primary sections captured via logarithm of total number of pre-primary students) for the $i$-th school from district $d$ of state $s$ measured at time $t$. For learning outcomes, $Y_{i\,dst}$ corresponds to the test score (math and reading score) of the $i$-th child from district $d$ of state $s$ at time $t$. The first outcome variable for enrolment is the availability of pre-primary section. It is a dummy that takes the value 0 if the $i$-th school does not have a pre-primary section and the value 1 if it has a pre-primary section. The second variable represents the number of pre-primary students enrolled in $i$-th school.

For learning outcomes, $Y_{i\,dst}$ takes integer values from 0 to 4, where 0 means no learning skills and 4 implies the highest level of learning. '$wb_s$' is a dummy that takes the value 1 if the i-th observation comes from West Bengal. All the standard errors are clustered at the district level. Even though the policy was implemented at the state level, the number of clusters is too small if we cluster the standard errors at the state level. To overcome this difficulty, we have also used the wild cluster bootstrap method, and the corresponding p values and standard errors are reported within parentheses in all the regression tables.

The '$post_t$' dummy compares the outcomes for the post-program years (2013 or after) when it assumes the value 1 to the same before 2013 (for '$post_t$'=0). The coefficient of interest is $\beta_2$ , which captures the intent-to-treat (ITT) estimate of the impact of the program. $X_{idst}$ includes all the child-, household-, school- and village-level characteristics. For school level regression on enrolment, $X_{idst}$ includes the total number of classrooms, the total number of teachers, and availability of electricity and playground in the school.

For child level regression on learning outcomes, $X_{idst}$ controls for the household variables for the i-th child that include the number of household members, presence of pucca house; presence of electricity; possession of a TV, and presence of a toilet. $X_{idst}$ also controls for child-level characteristics, including the child's age, whether the mother went to school, and village-level characteristics, such as presence of a government primary school, private school in the village, village post office, bank, and pucca road. We also include district ($\partial_{ds}$), year ($\gamma_t$) fixed effects and district-specific linear trend ($\partial_{ds} \times t$).

In the previous regression, the double difference compares the difference across the government schools in West Bengal and the control states before and after the policy. However, another way of constructing the DD specification is to compare changes between the treated government schools and the untreated private schools in West Bengal. To consider this alternative model, we introduce a '$govtschool_{id}$' dummy variable and estimate the following regression restricting the sample to only the schools of West Bengal.

$$(2)\quad y_{i\,dt} = \beta_0 + \beta_1\, govtschool_{id} + \beta_2\, post_t + \beta_3\, (govtschool_{id} \times post_t) + \vartheta X_{idt} + \gamma_t + \partial_d + (\partial_d \times t) + \epsilon_{i\,dt}$$

The dummy '$govtschool_{id}$' takes the value 1 if the *i*-th school is a public school run by the government and a value 0 if it is a private school. Similarly, for regressions on learning outcomes, '$govtschool_{id}$' takes the value 1 if the i-th child from district *d* goes to a government school and 0 otherwise.

The validity of the DD estimator relies on the parallel-trend assumption, which says that the outcome variable should move parallelly between the treated and control group in the absence of the treatment (Angrist and Pischke 2009). In the context of our study, if the trends of the outcome variables between the treated and control schools (children) evolved similarly before the program, they would continue to do so in the post-periods in the absence of the pre-primary program. To test the parallel trend assumption, we conducted year wise event study.

Essentially, the equations remain the same except that the '$post_t$' dummy is now replaced by continuous years. The event study graphs are presented in the appendix section.

*4.2 Triple difference*

Even though we get an estimate of the program using the DD method, certain broader trends can still affect the DD estimate. Therefore, a triple difference or difference-in-difference-in-differences approach (DDD) can be adopted to refine our results and eliminate any other trends that might be present. Similar to Muralidharan and Prakash (2017), we make use of a triple Difference (DDD) regression by comparing the DD estimates from West Bengal to a group of neighboring states, including Bihar, Orissa, and Jharkhand.

Once we have appropriately chosen the control states, we can estimate the program's impact by comparing the changes in outcomes between pre- and post-periods across government and private schools in West Bengal to the above-mentioned control states. We estimate the following triple-difference (DDD) model:

$$(3)\quad y_{idst} = \beta_0 + \beta_1.\ govtschool_{ids} + \beta_2.post_t + \beta_3.(wb_s \times govtschool_{ids} \times post_t)$$
$$+ \beta_4.(govtschool_{ids} \times\ post_t) + \beta_5.(\ post_t\ \times\ wb_s)$$
$$+ \beta_6.(govtschool_{ids}\ \times\ wb_s) + \vartheta X_{idst} + \gamma_t + \partial_{ds} + (\partial_{ds}\ \times\ t) + \epsilon_{i\,dst}$$

Here, the '$wb_s$' dummy takes the value 1 if the school in question is in the state of West Bengal and 0 otherwise. The rest of the variables used in equation (3) has already been discussed while discussing the DD model. In this model, $\beta_3$, the ITT estimate is our coefficient of interest.

*4.3 Synthetic control*

In our DD and DDD specifications, we selected the control states on the basis of geographical proximity and similarities in cultural and socioeconomic characteristics. However, there could still be pre-existing differences between these groups of states, making the comparison inaccurate. To rule out any such possibility, we use the synthetic control method (SCM) following the recent literature (Abadie and Gardeazabal 2003; Abadie, Diamond, and Hainmueller 2010; Peri and Yesenov, 2019). In our context, SCM can potentially be a superior method than the classic difference-in-differences because it takes a data-driven approach to obtain a linear combination of states to form a suitable control group for WB.[4] We used the ASER and DISE databases from 2009 to 2018 for this purpose. Using the ASER data, we calculated the state-level annual estimates of learning outcomes among government school children aged 5-6 years.

To obtain the synthetic control estimates for the pre-primary enrolment, we have used state-level annual estimates of the proportion of schools with pre-primary sections and the total number of students in pre-primary sections by using DISE data. The predictors of the pre-primary enrolment are the state-wise average number of classrooms in government schools, the proportion of government schools with electricity, and playgrounds. For learning outcomes, predictor variables for test scores are the proportion of households that are fully cemented, average household size, proportion of households having a TV, proportion of households with electricity, proportion of households with toilet, proportion of mothers who went to school at some point of time, proportion of villages with the pucca road, the proportion of villages with a bank, proportion of villages with post office, the proportion of villages with private school and proportion of villages with government primary school. The outcome variables from 2009 to 2012 are also used as predictors.[5]

[4] The control states for the synthetic control method are Himachal Pradesh, Punjab, Uttarakhand, Haryana, Rajasthan, Uttar Pradesh, Bihar, Sikkim, Nagaland, Manipur, Mizoram, Tripura, Meghalaya, Assam, Jharkhand, Orissa, Chhattisgarh, Madhya Pradesh, Gujarat, Maharashtra, Andhra Pradesh, Karnataka, Kerala and Tamil Nadu.
[5] The weights for the individual states and predictor balance between the treated and the synthetic group are reported in Appendix Tables 1-4 for both enrolment and learning outcomes.

## 5. Results

We present the results related to the impact of the pre-primary program on enrolment in Section 5.1 and then learning outcomes in Section 5.2.

### *5.1 The impact on enrolment*

We begin this section with the results from the regression Equation 1 in Panel A of Table 2. The coefficient $\beta_3$ associated with the interaction term between '$wb_s$' and '$post_t$' is positive and highly significant for both the dependent variables considered, with a 91 percentage points increase in the probability of a government school in WB having a pre-primary section after 2013 and ($e^{2.06}$-1) × 100 ≈ 684 percent increase in the difference in pre-primary enrolment between government schools in WB and its neighboring states after 2013.

**[Insert Table 2 here]**

Next, we investigated if the government schools in West Bengal are significantly different from the private schools in the state in terms of availability of pre-primary sections and enrolment in them. We estimated Equation 2 and presented the results in Panel B of Table 2. The coefficient $\beta_2$ associated with the interaction term between '$govtschool_{id}$' and '$post_t$' dummy is positive and highly significant for both the outcome variables. The result shows that the probability of having a pre-primary section in government schools vis-à-vis private schools in West Bengal after 2013 increased by more than 58 percentage points. On the other hand, the difference in pre-primary enrolment between government and private schools in the state shows ($e^{1.41}$-1) × 100 ≈ 309 percent increase after the program was introduced.

Finally, we employed the triple difference method by estimating Equation 3, and the findings from Panel C of Table 3 confirm a positive significant impact of the program on both the availability of and enrollment in the pre-primary section of government schools. The difference in the probability of having a pre-primary section in a government school vis-à-vis a private school after 2013 between West Bengal and its neighbouring states increased by 72 percentage points compared to the same difference before 2013. Similarly, the impact on pre-primary enrollment is estimated to be ($e^{1.77}$-1) × 100 ≈ 487 percent. So, the DD and DDD regression results suggest that the pre-primary program successfully increased enrolment and availability of pre-primary education in government schools.

*5.2 The impact on learning outcomes*

First, we consider children in the pre-primary age group in government schools of West Bengal and compare them to the children of the same age group from government schools in the control states. In doing so, we estimate equation 1 for learning outcomes. The results presented in columns 2 and 4 of Panel A in Table 3 suggest that the coefficient associated with the interaction term between '$wb_s$' and '$post_t$' is negative and significant for both the test scores. However, the event study graphs in Panels A and B in Figure A2 show that trends for 2011 and 2012 drive the decline in test scores. This implies that reading and math scores of government school pre-primary students in West Bengal were significantly higher than the same cohort of children in control states for 2011 and 2012 only. We cannot attribute the negative and significant coefficient to the impact of the program only.

**[Insert Table 3 here]**

Next, we bring in the private schools of West Bengal and investigate if the test scores of government school pre-primary children in the state are significantly different from those private school children. The coefficients associated with the interaction term between '$post_t$' and '$govtschool_{id}$' are found to be negative and statistically insignificant for both raw test scores in Panel B of Table 3. So, government school students of WB in the pre-primary age group do not perform significantly better in math and reading tests than children from the same category in private schools.

Finally, we consider the DDD model presented by equation 3. Panel C of Table 3 shows that the coefficient of the triple interaction term is statistically not significant. Thus, the major findings from the regressions on learning outcomes indicate that both math and reading scores did not improve among the government school children after the program's implementation.

*5.3 Results from synthetic control analysis*

In the synthetic control method (SCM), we compare the government schools in WB to those in its synthetic control counterpart. The synthetic control methodology results reaffirm our findings from the previous analyses using the first DD regression specification. First, we plot the trajectory of the dependent variables for both WB and its synthetic counterpart. Next, we take the difference in the outcome variable between WB and the synthetic control states and trace the path of this gap over the years before and after the program.

We also adopted a bias-corrected approach of the SCM as there can be bias in the estimated marginal treatment effects because of discrepancies between the values of the predictor variable in each treated unit and its synthetic control donors (Wiltshire 2021). Similar to the approach in Abadie and Imbens (2011) to address inexact matching on predictor variables, a bias-correction method has been proposed in the recent literature (Abadie and L'Hour 2021; Ben-Michael, Feller, and Rothstein 2021). We call this later one the 'Bias-corrected' estimates and the original one 'Classic' estimates (Abadie et al. 2010). We present the 'gaps' or the effect graphs with the 'Classic' and 'Bias-corrected' estimates.[6]

**[Insert Figure 3 here]**

Panels A and C of Figure 3 present the trends in the percentage of government schools with pre-primary and the number of pre-primary students in government schools, respectively. Panels B and D of Figure 3 show the corresponding gaps in the outcome variable between West

[6] These bias-corrected graphs are created using Stata's 'allsynth' command (Wiltshire 2021).

Bengal and the synthetic control state. Our estimate of the effect of the pre-primary program on enrolment and learning outcomes is the difference between the outcome variables (both enrolment and learning scores) in WB and its synthetic version after the pre-primary program in 2013. Immediately after the implementation, the two lines of pre-primary availability and pre-primary enrolment diverge noticeably. We see a significant increase in pre-primary enrolment (both availability of the pre-primary section and number of pre-primary students) following the program's implementation in WB compared to the synthetic state.

For reading and mathematics scores, we used the same set of predictors to generate a similar figure as above. Panels A and C of Figure 4 shows that both math and reading scores did not change significantly in WB after the program's implementation.

**[Insert Figure 4 here]**

The dotted line corresponds to the synthetic state. As we can see the lines of math and reading score for WB do not significantly diverge from the synthetic state, The gap between the two lines, essentially gives us the effect of the program. Panels B and D in Figure 4 demonstrate that the pre-primary program did not impact math and learning scores as it is very close to the zero mark.

We also conducted a series of placebo run by iteration. In each iteration, we choose one of the control states and reassign the program intervention to that state. By doing so, we shift WB to the donor pool. We estimate the impact associated with each placebo. For each control state, we obtain different values of the impact, eventually leading to a distribution of estimated impacts for the donor pool. The figures for the placebo run are given in the appendix.

## 6. Potential Mechanism for Null Impact on Learning Outcome: School Infrastructure

Our findings show that the program does not have any impact on the test scores. We hypothesize that the null impact may be due to structural factors like school facilities and the availability of teachers. For instance, increased enrolment without a commensurate improvement in the availability of facilities and teachers might result in congestion externality and have an adverse effect on academic skills.

Using the DISE data, first, we compared indicators of school infrastructure in government schools in WB to the same category of schools in the control states. We considered indicators like the average number of classrooms, the average number of teachers in the pre-primary section and the total number of teachers. Next, we compared the government schools in WB to the private schools in the state. Finally, the differences between government and private schools in WB are compared to the same difference in the control states.

**[Insert Table 4]**

Correspondingly, we used all three specifications for DD and DDD. The outcome variables were the logarithm of the total number of classrooms, the logarithm of number of teachers in pre-primary, and the logarithm of the total number of teachers. The result from Panel A of Table 4 shows that the difference in the number of classrooms in government schools between WB and the control states after the program's introduction declined by $(e^{.09}-1) \times 100 \approx 9$ percent compared to the same difference before the program was introduced. The most important result is the comparison between columns 2 and 3 in Panel A. Even though the difference in the number of pre-primary teachers in government schools between West Bengal and the control states increased by $(e^{.18}-1) \times 100 \approx 20$ percent after the introduction of the program, the difference in the number of total teachers decreased by $(e^{.04}-1) \times 100 \approx 4$ percent during the same time. It implies that the government schools in WB used the existing pool of teachers to teach at pre-primary sections without recruiting additional teachers. In fact, the growth in teacher recruitment for the government schools in the state was much less than the same category of schools in the control states. Similar stories come up from Panels B and C of Table 4. The difference in the number of total classrooms and the difference in the number of total teachers between the government and private schools in WB declined by ($e^{.12}$-1) × 100 ≈ 13 percent and ($e^{.13}$-1) × 100 ≈ 14 percent, respectively after the program was introduced. The triple difference estimate from Panel C also reports a decline of ($e^{.47}$-1) × 100 ≈ 60 percent in the difference in the number of total teachers between government and private schools in WB and the control states after 2013, whereas the difference in the number of pre-primary teachers across the same group increased by ($e^{.29}$-1) × 100 ≈ 34 percent.

We also adopted the synthetic control methodology again to rule out any bias in our results. Results from the synthetic control methodology are presented in Figure 5.

**[Insert Figure 5]**

It is evident from Figure 5 (Panel A-F) that infrastructure has deteriorated more in the government schools of WB after the introduction of pre-primary. The result from synthetic control reaffirms that although enrolment has increased, the average number of teachers in government schools is less than its synthetic counterpart. The result is in tune with the ground reality where the pre-primary students sit with students from higher standards. They do not receive any extra attention from the teachers. As a result, they fail to develop any learning skills even though they enjoy one extra year of schooling in pre-primary before being promoted to class 1.

## 7. Conclusion

Our estimates of the program indicate that it has a massive role in increasing pre-primary enrolment. Hence, the program has successfully sent children at their nascent age to schools.

While this effect is prevalent across the state, we also observe that the program fails to have an impact on the test scores of children.

Despite the positive influence on enrolment, we found no improvement in learning ability (both math as well as reading score). This phenomenon can be attributed to the declining school infrastructure that has taken place in West Bengal over the years after the program was announced. The announcements made by the state have not been reflected in the improvement of school infrastructure. Extant literature shows the importance of learning environment where classroom overcrowding can negatively affect student achievement by hampering the teaching and learning process (Angrist and Lavy 1999; Case and Deaton 1999). The decline in the availability of classrooms and teachers potentially accounts for the non-significant impact on learning outcomes in government schools.

Suppose the sole objective of the pre-primary program was to make children familiar with the school environment and make it a habit for them to attend classes. Our findings strongly suggest that the program has indeed been quite successful in increasing pre-primary enrolment. From the labor market perspective, a pre-primary program can also free up mothers' time from childcare activities. Thus, such a program can have a positive spillover effect through a higher female labour force participation rate (Halim, Johnson, and Perova 2022). However, if the benefits of pre-primary program in the form of better learning outcomes are to be realized, further steps such as heavy investments in school infrastructure need to be undertaken.

## References

Abadie, A., and Gardeazabal, J. 2003. "The economic costs of conflict: A case study of the Basque Country." *American economic review*, *93*(1), 113-132.

Abadie, A., Diamond, A., and Hainmueller, J. 2010. "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program." *Journal of the American statistical Association*, *105*(490), 493-505.

Abadie, A., and Imbens, G. W. 2011. "Bias-corrected matching estimators for average treatment effects." *Journal of Business & Economic Statistics*, *29*(1), 1-11.

Abadie, A., and L'Hour, J. 2021. "A penalized synthetic control estimator for disaggregated data." *Journal of the American Statistical Association*, *116*(536), 1817-1834.

Alcott, B., and Rose, P. 2017. "Learning in India's primary schools: How do disparities widen across the grades?" *International Journal of Educational Development*, *56*, 42-51.

Angrist, J. D., and Lavy, V. 1999. "Using Maimonides' rule to estimate the effect of class size on scholastic achievement." *The Quarterly journal of economics*, *114*(2), 533-575.

Angrist, J. D., and Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Annual Status of Education Report (Rural). 2018. Retrieved from http://img.asercentre.org/docs/ASER%202018/Release%20Material/aserreport2018.pdf.

Baker, M. 2011. "Innis Lecture: Universal early childhood interventions: what is the evidence base?" *Canadian Journal of Economics/Revue canadienne d'économique*, *44*(4), 1069-1105.

Behrman, J. R., and Urzúa, S. 2013. "Economic perspectives on some important dimensions of early childhood development in developing countries." In *Handbook of early childhood* *development research and its impact on global policy*, ed Britto, P. R., Engle, P. L., and Super, C. M. Oxford University Press.

Ben-Michael, E., Feller, A., & Rothstein, J. 2021. “The augmented synthetic control method.” *Journal of the American Statistical Association*, *116*(536), 1789-1803.

Berlinski, S., Galiani, S., and Manacorda, M. (2008). “Giving children a better start: Preschool attendance and school-age profiles.” *Journal of public Economics*, *92*(5-6), 1416-1440.

Berlinski, S., Galiani, S., and Gertler, P. 2009. “The effect of pre-primary education on primary school performance.” *Journal of public Economics*, *93*(1-2), 219-234.

Britto, P. R., Lye, S. J., Proulx, K., Yousafzai, A. K., Matthews, S. G., Vaivada, T., ... & Lancet Early Childhood Development Series Steering Committee. 2017. “Nurturing care: promoting early childhood development.” *The Lancet*, *389*(10064), 91-102.

Cascio, E. U. 2015. “The promises and pitfalls of universal early education.” *IZA World of Labor*.

Case, A., and Deaton, A. 1999. “School inputs and educational outcomes in South Africa.” *The Quarterly Journal of Economics*, *114*(3), 1047-1084.

Chakraborty, T., and Jayaraman, R. 2019. “School feeding and learning achievement: evidence from India's midday meal program.” *Journal of Development Economics*, *139*, 249-265.

Cunha, F., and Heckman, J. 2007. “The technology of skill formation.” *American economic review*, *97*(2), 31-47.

Das, U., and Sarkhel, P. 2023. “Does more schooling imply improved learning? Evidence from the Kanyashree Prakalpa in India.” *Economics of Education Review*, *94*, 102406.

Elango, S., García, J. L., Heckman, J. J., & Hojman, A. 2015. "Early childhood education." In *Economics of Means-Tested Transfer Programs in the United States, Volume 2* (pp. 235-297). University of Chicago Press.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., ... and Grantham-McGregor, S. 2014. "Labor market returns to an early childhood stimulation intervention in Jamaica." *Science*, *344*(6187), 998-1001.

García, J. L., Heckman, J. J., Leaf, D. E., and Prados, M. J. 2017. "The life-cycle benefits of an influential early childhood program (No. w22993)." National Bureau of Economic Research.

Halim, D., Johnson, H. C., and Perova, E. 2022. Preschool Availability and Women's Employment: Evidence from Indonesia. *Economic Development and Cultural Change*, *71*(1), 39-61.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., and Yavitz, A. 2010. "The rate of return to the HighScope Perry Preschool Program." *Journal of public Economics*, *94*(1-2), 114-128.

Heckman, J. J. 2011. "The economics of inequality: The value of early childhood education." *American Educator*, *35*(1), 31.

Heckman, J. J. 2013. "*Giving kids a fair chance*." Mit Press.

Lahoti, R., and Sahoo, S. 2020. "Are educated leaders good for education? Evidence from India." *Journal of Economic Behavior & Organization, 176*, 42-62.

Muralidharan, K., and Prakash, N. 2017. "Cycling to school: Increasing secondary school enrollment for girls in India." *American Economic Journal: Applied Economics*, *9*(3), 321-50.

NCERT (2017). “National Report to inform Policy, Practices and Teaching Learning: National Achievement Survey - NAS (2017) Classes III, V and VIII.” New Delhi, NCERT and MHRD. Retrieved from https://ncert.nic.in/pdf/NAS/WithReleaseDate_NPPTL.pdf on 17 May 2023; 15:10 hrs

NEP 2020. Policy document released by Government of India. Retrieved from https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English.pdf on 17 May 2023; 14.35 hrs

Nores, M., and Barnett, W. S. 2010. “Benefits of early childhood interventions across the world:(Under) Investing in the very young.” *Economics of education review*, *29*(2), 271-282.

Peri, G., & Yasenov, V. 2019. “The labor market effects of a refugee wave synthetic control method meets the mariel boatlift.” *Journal of Human Resources*, *54*(2), 267-309.

Sayre, R. K., Devercelli, A. E., Neuman, M. J., & Wodon, Q. 2015. “Investing in early childhood development: review of the World Bank's recent experience.”

UNICEF. 2018. “Catalysing transformational change in school education”

UNICEF. 2019. “A World Ready to Learn: Prioritizing Quality Early Childhood Education.” New York: UNICEF.

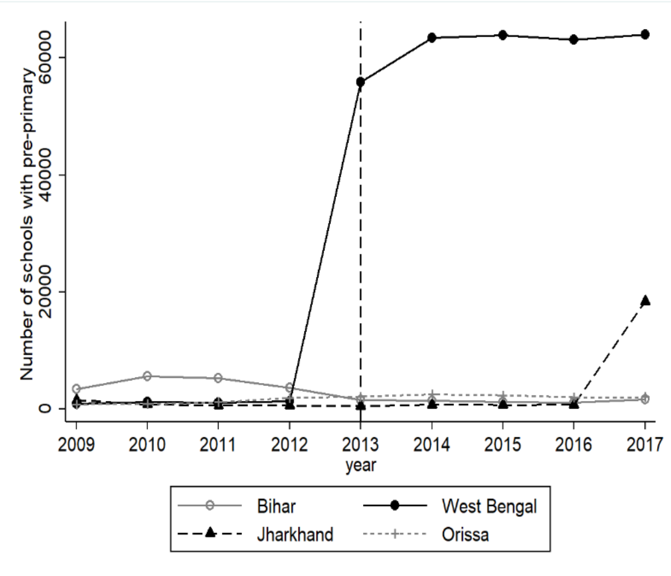
United Nations General Assembly. 2015. “Transforming our world: the 2030 Agenda for Sustainable Development.” *United Nations: New York, NY, USA.*

Wiltshire, J. C. 2021. “allsynth: Synthetic Control Bias-Correction Utilities for Stata.” In *In2021 Stata Conference* (No. 15).

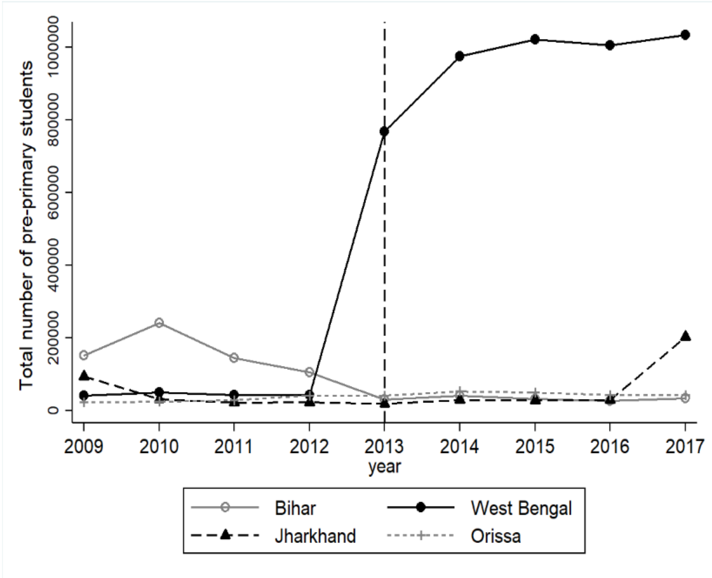
Wotipka, C. M., Rabling, B. J., Sugawara, M., and Tongliemnak, P. 2017. “The worldwide expansion of early childhood care and education, 1985–2010.” *American Journal of Education*, *123*(2), 000-000.

## Figure 1: Trends in enrollment

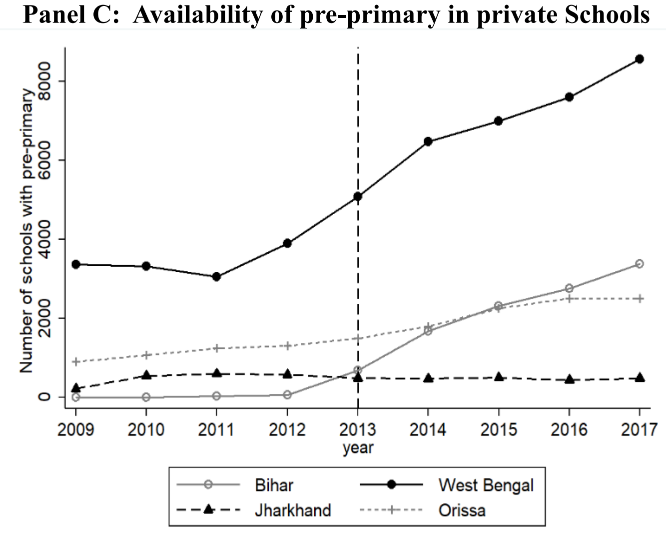
**Panel A: Availability of pre-primary in government schools**



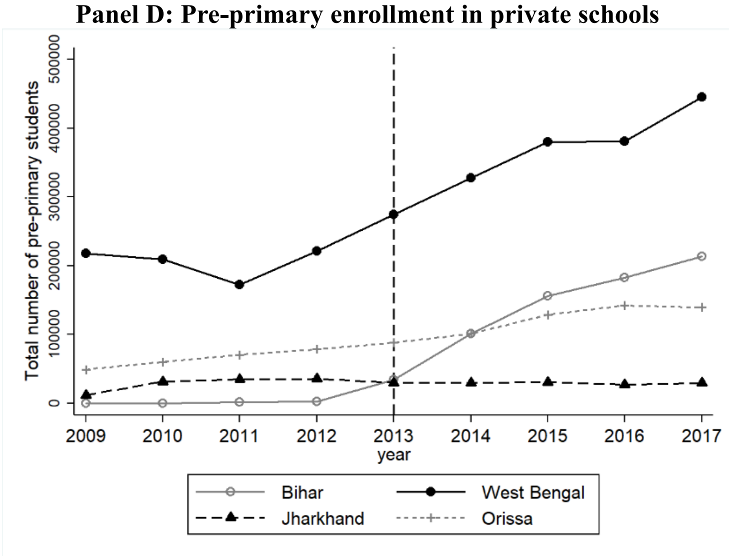

**Panel B: Pre-primary enrollment in government schools**




**Panel C: Availability of pre-primary in private Schools**

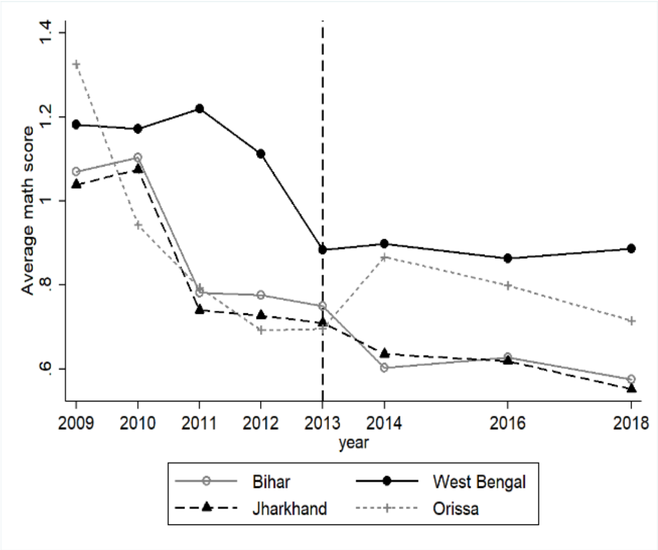



**Panel D: Pre-primary enrollment in private schools**

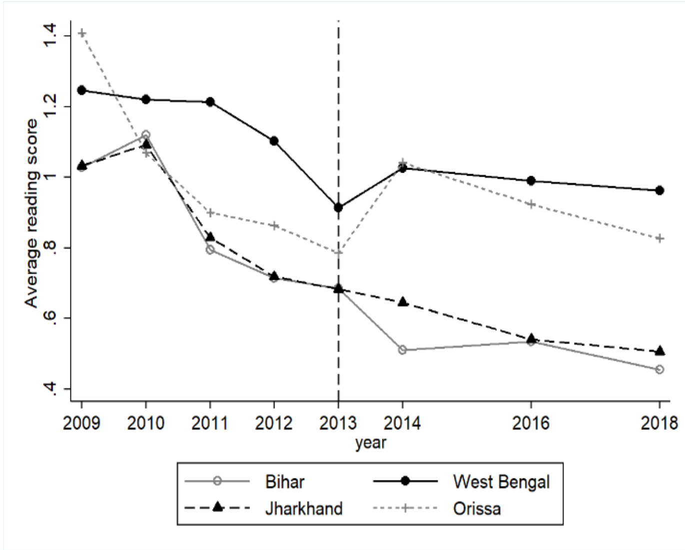


## Figure 2: Trends in learning outcomes

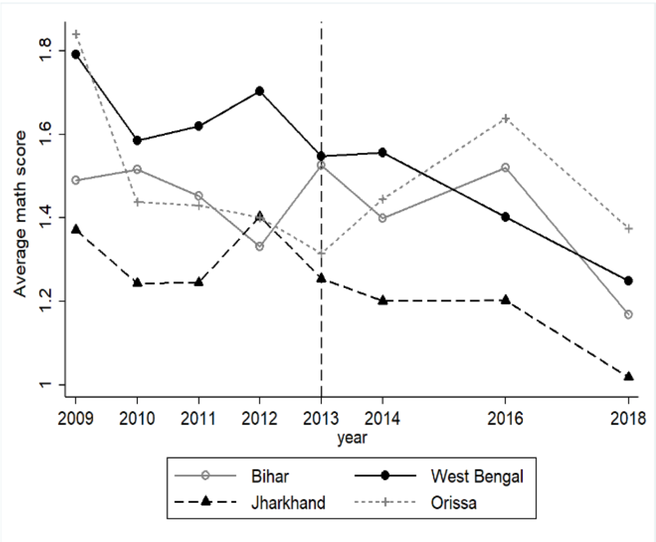
**Panel A: Average math score of 5-6 years old government school students**



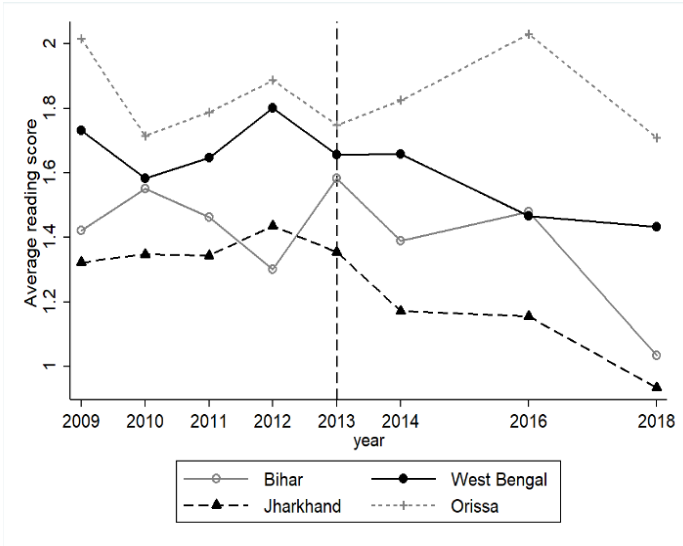

**Panel B: Average reading score of 5-6 years old government school students**




**Panel C: Average math score of 5-6 years old private school students**

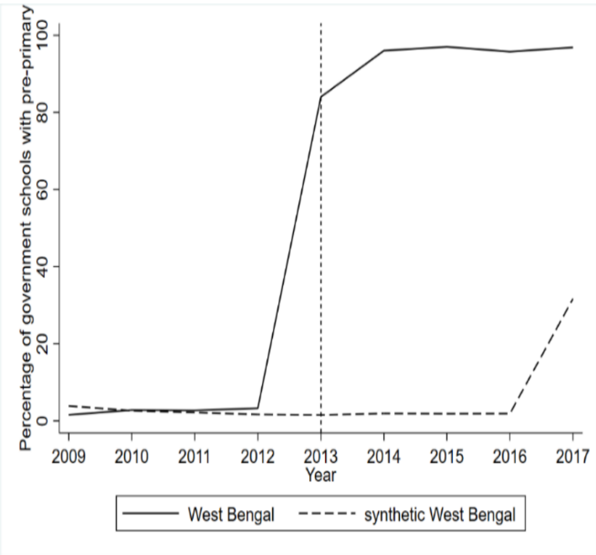



**Panel D: Average reading score of 5-6 years old government school students**

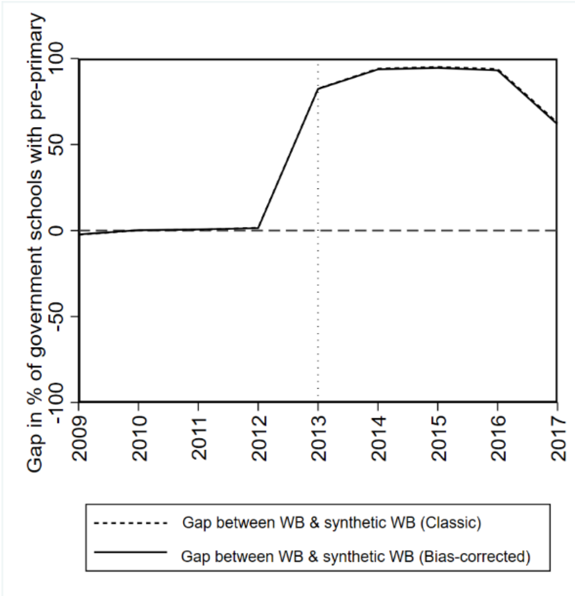


## Figure 3: Synthetic control analysis for enrollment

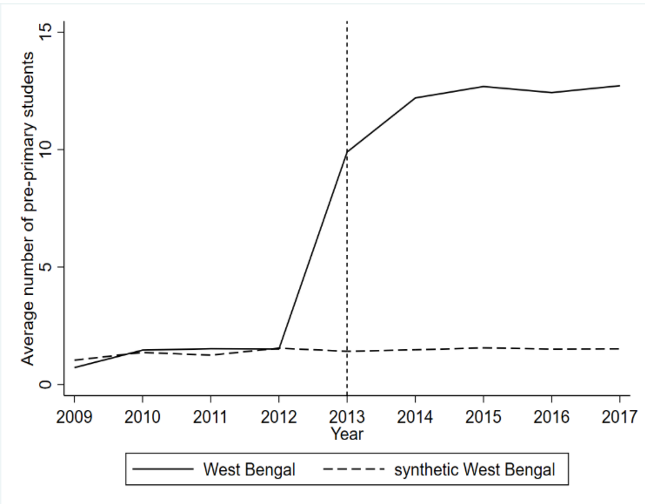
**Panel A: Availability of pre-primary in government schools of West Bengal and synthetic control**



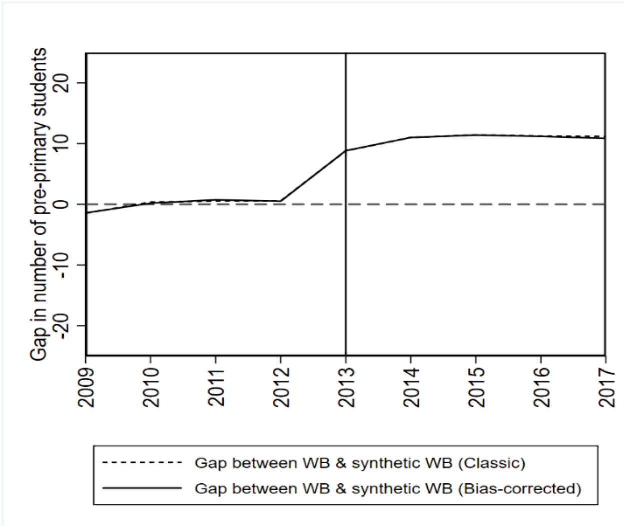

**Panel B: Gap between West Bengal and synthetic control in the availability of pre-primary in government schools**




**Panel C: Pre-primary enrolment in government schools of West Bengal and synthetic control**

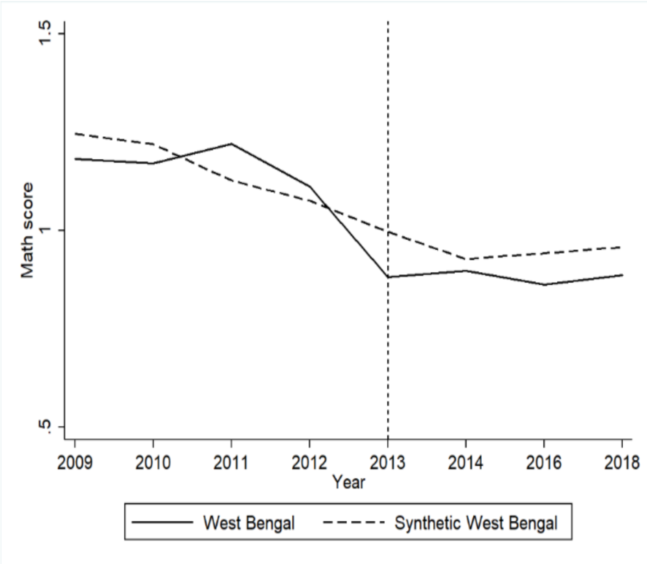



**Panel D: Gap between West Bengal and synthetic control in pre-primary enrolment in government schools**

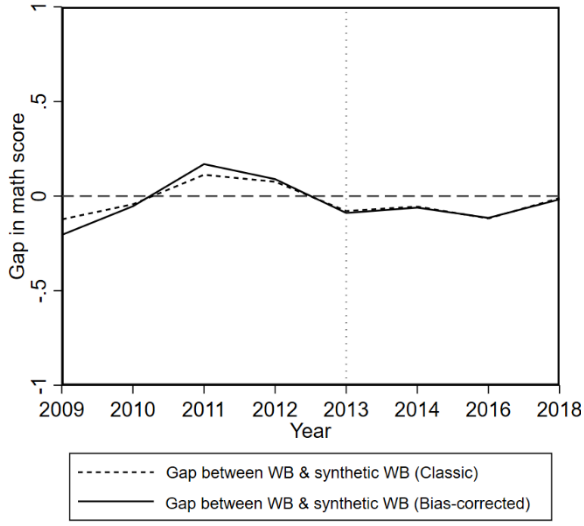


## Figure 4: Synthetic control analysis for learning outcomes

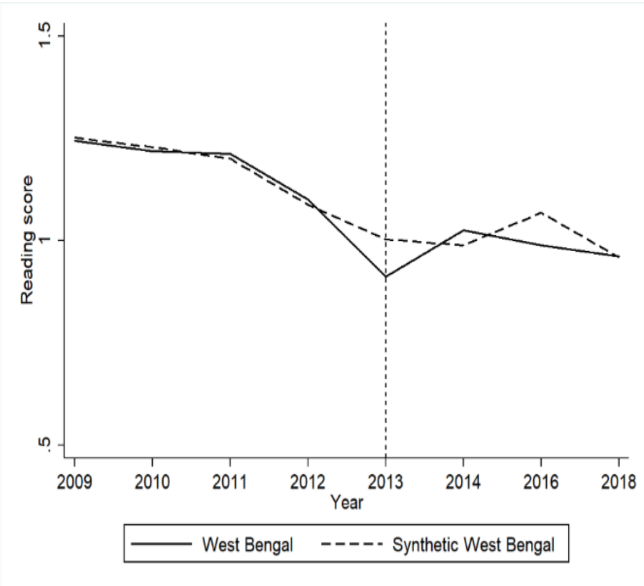
**Panel A: Math score of 5-6 years old in government schools of West Bengal and synthetic control**



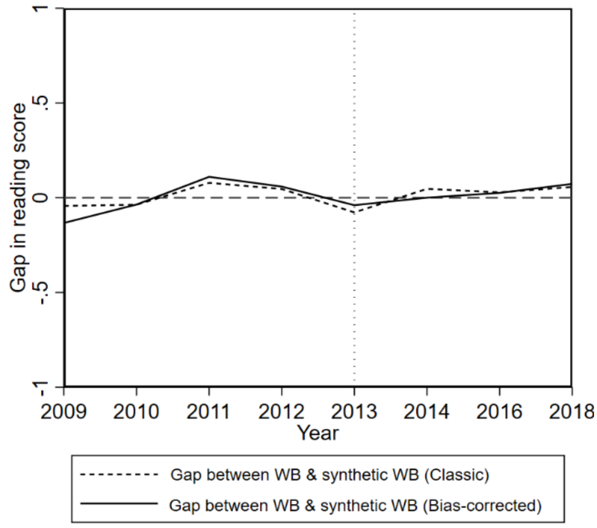

**Panel B: Gap in math score of 5-6 years old in government schools of West Bengal and synthetic control**




**Panel C: Reading score of 5-6 years old in government schools of West Bengal and synthetic control**

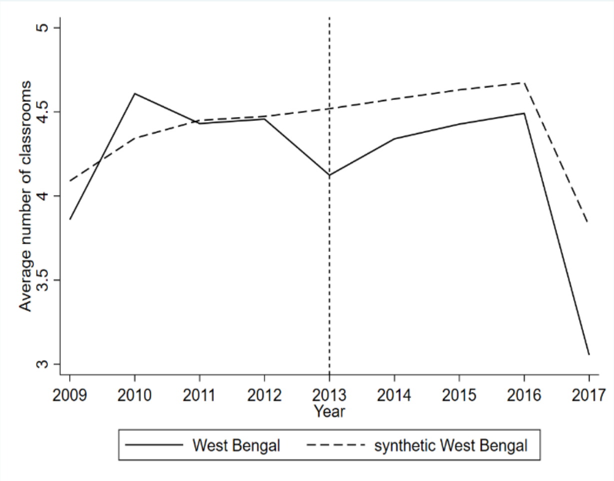



**Panel D: Gap in reading score of 5-6 years old in government schools of West Bengal and synthetic control**

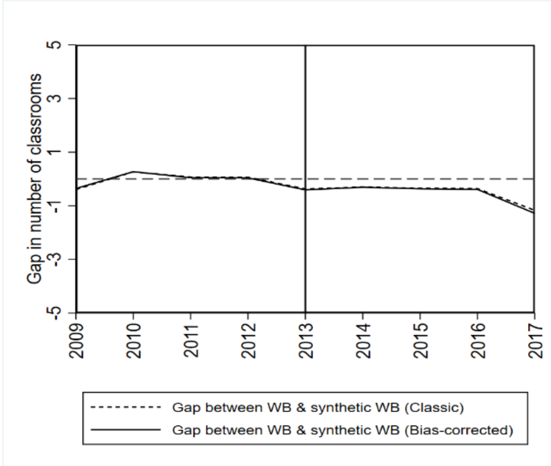


## Figure 5: Synthetic control analysis for school infrastructure

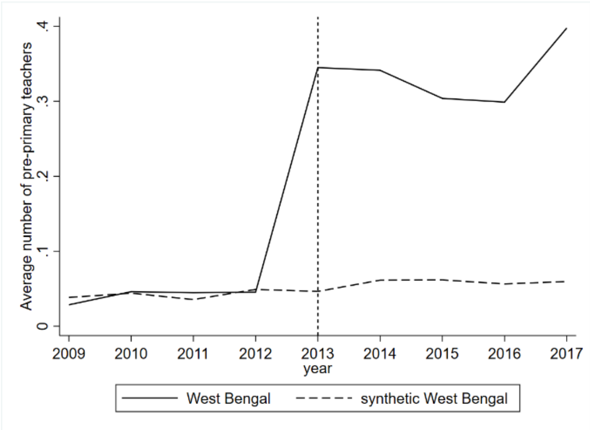
**Panel A: Average number of classrooms in government schools with pre-primary**



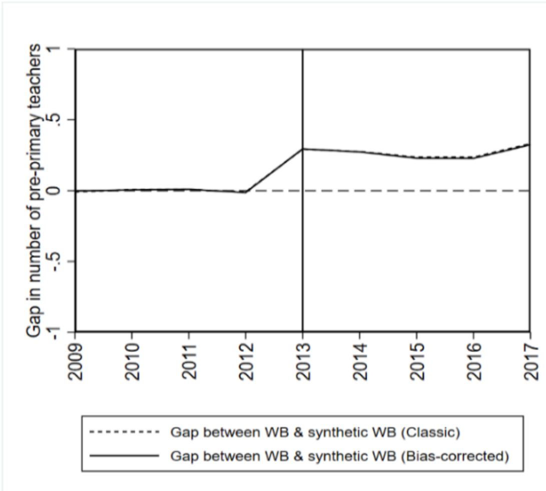

**Panel B: Gap in number of classrooms in government schools with pre-primary between West Bengal and synthetic control**



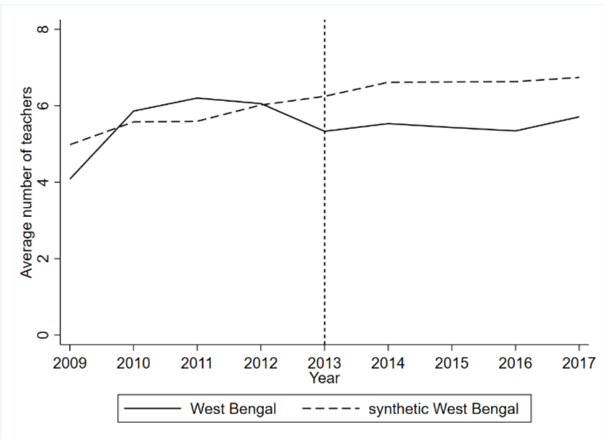

**Panel C: Average number of pre-primary teachers in government schools**



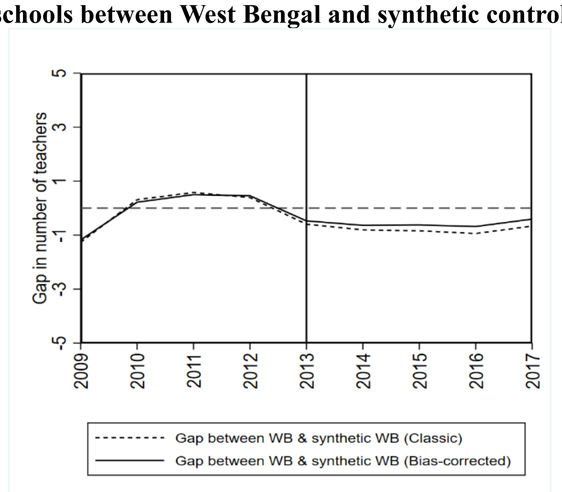

**Panel D: Gap in number of pre-primary teachers in government schools between West Bengal and synthetic control**




**Panel E: Average number of teachers in government schools**




**Panel F: Gap in number of teachers in government schools between West Bengal and synthetic control**



## Table 1: Summary Statistics

| | West Bengal | | | | | | Control states | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Government schools | | | Private schools | | | Government schools | | | Private schools | | |
| | N | Mean | SD | N | Mean | SD | N | Mean | SD | N | Mean | SD |
| ***School-level characteristics*** | | | | | | | | | | | | |
| Availability of pre-primary | 541,801 | 0.58 | 0.49 | 75,637 | 0.64 | 0.48 | 1,447,223 | 0.04 | 0.21 | 42,871 | 0.70 | 0.46 |
| Number of students in pre-primary | 541,801 | 9.19 | 13.46 | 75,637 | 34.77 | 46.20 | 1,447,223 | 1.12 | 8.82 | 42,871 | 42.23 | 50.11 |
| Availability of electricity | 541,801 | 0.47 | 0.50 | 75,637 | 0.62 | 0.49 | 1,447,223 | 0.16 | 0.36 | 42,871 | 0.75 | 0.43 |
| Availability of playground | 541,801 | 0.34 | 0.48 | 75,637 | 0.39 | 0.49 | 1,447,223 | 0.29 | 0.45 | 42,871 | 0.66 | 0.47 |
| Total classrooms in the school | 541,801 | 3.51 | 1.63 | 75,637 | 5.50 | 3.97 | 1,447,223 | 4.08 | 2.98 | 42,871 | 9.46 | 6.38 |
| Total teacher in the school | 541,801 | 3.58 | 1.84 | 75,637 | 7.24 | 5.13 | 1,447,223 | 4.07 | 3.14 | 42,871 | 10.44 | 9.42 |
| Number of pre-primary teachers | 541,801 | 0.26 | 0.83 | 75,637 | 1.80 | 2.60 | 1,447,223 | 0.05 | 0.39 | 42,871 | 2.36 | 2.76 |
| ***Child learning outcomes (5-6 years)*** | | | | | | | | | | | | |
| Math score | 6,550 | 1.03 | 0.91 | 1,755 | 1.52 | 0.94 | 55,863 | 0.81 | 0.94 | 8,108 | 1.33 | 1.05 |
| Reading score | 6,550 | 1.10 | 1.06 | 1,755 | 1.60 | 1.09 | 55,863 | 0.80 | 1.03 | 8,108 | 1.39 | 1.27 |
| Gender | 6,550 | 0.50 | 0.50 | 1,755 | 0.55 | 0.50 | 55,863 | 0.53 | 0.50 | 8,108 | 0.60 | 0.49 |
| ***Household-level characteristics*** | | | | | | | | | | | | |
| Household pucca or not | 6,550 | 0.20 | 0.40 | 1,755 | 0.42 | 0.49 | 55,863 | 0.22 | 0.42 | 8,108 | 0.50 | 0.50 |
| Household size | 6,550 | 5.94 | 2.65 | 1,755 | 6.03 | 2.98 | 55,863 | 6.84 | 3.00 | 8,108 | 7.21 | 3.50 |
| Availability of toilet | 6,550 | 0.47 | 0.50 | 1,755 | 0.77 | 0.42 | 55,863 | 0.20 | 0.40 | 8,108 | 0.50 | 0.50 |
| Availability of electricity | 6,550 | 0.72 | 0.45 | 1,755 | 0.88 | 0.32 | 55,863 | 0.53 | 0.50 | 8,108 | 0.79 | 0.40 |
| Possession of TV | 6,550 | 0.39 | 0.49 | 1,755 | 0.70 | 0.46 | 55,863 | 0.20 | 0.40 | 8,108 | 0.51 | 0.50 |
| Mother went to school | 6,550 | 0.66 | 0.47 | 1,755 | 0.86 | 0.34 | 55,863 | 0.42 | 0.49 | 8,108 | 0.74 | 0.44 |
| ***Village-level characteristics*** | | | | | | | | | | | | |
| Availability of pucca road | 6,550 | 0.58 | 0.49 | 1,755 | 0.74 | 0.46 | 55,863 | 0.67 | 0.47 | 8,108 | 0.79 | 0.40 |
| Availability of post office | 6,550 | 0.36 | 0.48 | 1,755 | 0.48 | 0.50 | 55,863 | 0.33 | 0.47 | 8,108 | 0.41 | 0.49 |
| Availability of bank | 6,550 | 0.24 | 0.42 | 1,755 | 0.32 | 0.47 | 55,863 | 0.18 | 0.39 | 8,108 | 0.32 | 0.47 |
| Availability of government school | 6,550 | 0.95 | 0.21 | 1,755 | 0.94 | 0.23 | 55,863 | 0.92 | 0.26 | 8,108 | 0.93 | 0.25 |
| Availability of private school | 6,550 | 0.27 | 0.44 | 1,755 | 0.53 | 0.50 | 55,863 | 0.26 | 0.44 | 8,108 | 0.47 | 0.50 |

Notes: 'N' denotes the number of observations in each category and 'SD' represents the corresponding standard deviation of that group.

## Table 2: DD and DDD results for availability of pre-primary and enrolment in pre-primary level

| | (1)<br>Availability of<br>pre-primary schools | (2)<br>Log (1+number of<br>pre-primary students) |
|---|---|---|
| **Panel A: DD Specification 1** | | |
| West Bengal x Post | **0.91*** ** | **2.06*** ** |
| | **(0.04)** | **(0.11)** |
| | (p val=0.02, ci[0.54, 1.43]) | (p val=0.06, ci[-1.04, 4.23]) |
| School-level controls | Yes | Yes |
| Year fixed effect | Yes | Yes |
| District fixed effect | Yes | Yes |
| District-specific linear trend | Yes | Yes |
| No of Observations | 1,989,024 | 1,989,024 |
| R Squared | 0.75 | 0.65 |
| | | |
| **Panel B: DD Specification 2** | | |
| Government school | -0.38*** | -1.16*** |
| | (0.05) | (0.17) |
| Government school x Post | **0.58*** ** | **1.41*** ** |
| | **(0.05)** | **(0.18)** |
| | (p val=0.00, ci [0.44, 0.70]) | (p val=0.00, ci [0.96, 1.80]) |
| School-level controls | Yes | Yes |
| Year fixed effect | Yes | Yes |
| District fixed effect | Yes | Yes |
| District-specific linear trend | Yes | Yes |
| No of Observations | 617,438 | 617,438 |
| R Squared | 0.78 | 0.58 |
| | | |
| **Panel C: DDD Specification** | | |
| Government school | -0.52*** | -1.74*** |
| | (0.02) | (0.11) |
| West Bengal x Post | 0.17*** | 0.27 |
| | (0.05) | (0.18) |
| Post x Government school | -0.15*** | -0.46*** |
| | (0.02) | (0.10) |
| Government school x West Bengal | 0.12** | 0.38* |
| | (0.05) | (0.21) |
| West Bengal x Government school x Post | **0.72*** ** | **1.77*** ** |
| | **(0.06)** | **(0.21)** |
| | (p val=0.06, ci [-0.07, 1.10]) | (p val=0.08, ci[-1.28, 4.59]) |
| School level controls | Yes | Yes |
| Year fixed effect | Yes | Yes |
| District fixed effect | Yes | Yes |
| District-specific linear trend | Yes | Yes |
| No of Observations | 2,107,532 | 2,107,532 |
| R Squared | 0.72 | 0.59 |

Notes: The school-level control variables include the total number of classrooms, the total number of teachers, and the availability of electricity and playground in the school. The difference in number of observations across Panels A, B and C is due to the choice of control group. In Panel A, only the government schools of WB and the control states have been considered. In Panel B, only the government and private schools of WB have been considered. In Panel C, the government as well as the private schools are considered for both WB and the control states. Standard errors are clustered at the district level. ***=significant at 1% level, **=significant at 5% level, *=significant at 10% level. p-value and standard errors using the wild cluster bootstrap method are given in the parenthesis. Source: DISE 2009–17; authors' own calculations.

## Table 3: DD and DDD results for learning outcomes

| | Math score | | Reading score | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **Panel A: DD Specification 1** | | | | |
| West Bengal x Post | -0.24*** | -0.26*** | -0.11 | -0.13** |
| | (0.05) | (0.05) | (0.07) | (0.06) |
| | (p val=0.64., ci [-7.68,7.16]) | (p val=0.62, ci [-8.83,5.05]) | (p val=0.56, ci [-7.55,9.13]) | (p val=0.56, ci [-10.22,10.48]) |
| Household-level controls | No | Yes | No | Yes |
| Child-level controls | No | Yes | No | Yes |
| Village-level controls | No | Yes | No | Yes |
| Year fixed effect | Yes | Yes | Yes | Yes |
| District fixed effect | Yes | Yes | Yes | Yes |
| District-specific linear trend | Yes | Yes | Yes | Yes |
| No of Observations | 62,413 | 62,413 | 62,413 | 62,413 |
| R Squared | 0.09 | 0.13 | 0.10 | 0.14 |
| **Panel B: DD Specification 2** | | | | |
| Government school | -0.48*** | -0.32*** | -0.49*** | -0.31*** |
| | (0.05) | (0.05) | (0.06) | (0.06) |
| Government school x Post | -0.04 | -0.04 | -0.10 | -0.10 |
| | (0.04) | (0.04) | (0.06) | (0.07) |
| | (p val=0.40, ci [-0.12, 0.05]) | (p val=0.36, ci [-0.12, 0.05]) | (p val=0.13, ci [-0.23, 0.03]) | (p val=0.15, ci [-0.26, 0.05]) |
| Household-level controls | No | Yes | No | Yes |
| Child-level controls | No | Yes | No | Yes |
| Village-level controls | No | Yes | No | Yes |
| Year fixed effect | Yes | Yes | Yes | Yes |
| District fixed effect | Yes | Yes | Yes | Yes |
| District-specific linear trend | Yes | Yes | Yes | Yes |
| No of Observations | 8,305 | 8,305 | 8,305 | 8,305 |
| R Squared | 0.13 | 0.21 | 0.10 | 0.19 |
| **Panel C: DDD Specification** | | | | |
| Government school | -0.49*** | -0.36*** | -0.55*** | -0.40*** |
| | (0.03) | (0.03) | (0.04) | (0.03) |
| West Bengal x Post | -0.24*** | -0.26*** | -0.10 | -0.13 |
| | (0.07) | (0.06) | (0.09) | (0.09) |
| Post x Government school | -0.11*** | -0.10*** | -0.13*** | -0.11*** |
| | (0.04) | (0.03) | (0.04) | (0.04) |
| Government school x West Bengal | 0.02 | 0.01 | 0.06 | 0.06 |
| | (0.06) | (0.05) | (0.06) | (0.06) |
| West Bengal x Government school x Post | 0.07 | 0.05 | 0.03 | 0.01 |
| | (0.05) | (0.05) | (0.07) | (0.08) |
| | (p val=0.59, ci [-1.93,2.17]) | (p val=0.61, ci [-1.60,1.35]) | (p val=0.50, ci [-1.66,3.13]) | (p val=0.45, ci [-2.06,1.70]) |
| Household-level controls | No | Yes | No | Yes |
| Child-level controls | No | Yes | No | Yes |
| Village-level controls | No | Yes | No | Yes |
| Year fixed effect | Yes | Yes | Yes | Yes |
| District fixed effect | Yes | Yes | Yes | Yes |
| District-specific linear trend | Yes | Yes | Yes | Yes |
| No of Observations | 72,276 | 72,276 | 72,276 | 72,276 |
| R Squared | 0.11 | 0.16 | 0.12 | 0.17 |

*Note*: Household controls include the number of household members, whether the house was pucca or not; whether it had electricity; possession of a TV, presence of a toilet. Child-level characteristics include the child's age, whether the mother went to school. Village-level factors controlled for include whether the village has a government primary school, private school, village post office, bank, pucca road in the village. linear regression coefficients are presented with clustered standard errors at the district level given in parenthesis. Children aged 5 and 6 years have been considered for analysis. In Panel A, only the government school children of WB and the control states have been considered. In Panel B, only the government and private school children of WB have been considered. In Panel C, the government as well as the private school children are considered for both WB and the control states. ***=significant at 1% level, **=significant at 5% level, *=significant at 10% level. p-value and standard errors using the wild cluster bootstrap method are given in the parenthesis. Source: ASER 2009–18; authors' own calculations

## Table 4: DD and DDD results for school infrastructure

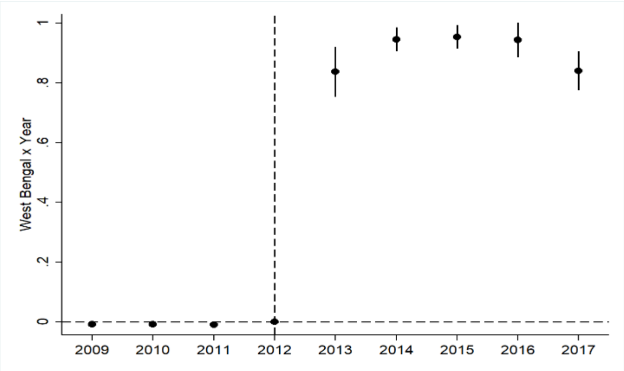
| | (1) Log (1+Total number of classrooms) | (2) Log (1+number of pre-primary teachers) | (3) Log (1+number of total teachers) |
|---|---|---|---|
| **Panel A: DD Specification 1** | | | |
| West Bengal x Post | **-0.09*** ** | **0.18*** ** | **-0.04*** |
| | **(0.02)** | **(0.04)** | **(0.02)** |
| | (p val=0.19, ci [-0.39,0.85]) | (p val=0.06., ci [-0.28,0.67]) | (p val=0.28., ci [-0.13,0.18]) |
| School-level controls | Yes | Yes | Yes |
| Year fixed effect | Yes | Yes | Yes |
| District fixed effect | Yes | Yes | Yes |
| District-specific linear trend | Yes | Yes | Yes |
| No of Observations | 1,989,024 | 1,989,024 | 1,989,024 |
| R Squared | 0.37 | 0.18 | 0.47 |
| **Panel B: DD Specification 2** | | | |
| Government school | 0.09*** | -0.46*** | -0.22*** |
| | (0.03) | (0.06) | (0.04) |
| Government school x Post | **-0.12*** ** | 0.07 | **-0.13*** ** |
| | **(0.02)** | (0.07) | **(0.03)** |
| | (p val=0.00, ci [-0.17,-0.08]) | (p val=0.37, ci [-0.08,0.24]) | (p val=0.68, ci [-0.02,0.03]) |
| School-level controls | Yes | Yes | Yes |
| Year fixed effect | Yes | Yes | Yes |
| District fixed effect | Yes | Yes | Yes |
| District-specific linear trend | Yes | Yes | Yes |
| No of Observations | 617,438 | 617,438 | 617,438 |
| R Squared | 0.33 | 0.28 | 0.42 |
| **Panel C: DDD Specification** | | | |
| Government school | -0.01 | -0.62*** | -0.22*** |
| | (0.05) | (0.03) | (0.07) |
| West Bengal x Post | -0.03 | -0.11* | 0.42*** |
| | (0.07) | (0.06) | (0.13) |
| Post x Government school | -0.03 | -0.26*** | 0.32*** |
| | (0.07) | (0.03) | (0.12) |
| Government school x West Bengal | 0.14*** | 0.09 | 0.03 |
| | (0.05) | (0.07) | (0.08) |
| West Bengal x Government school x Post | -0.07 | **0.29*** ** | **-0.47*** ** |
| | (0.07) | **(0.07)** | **(0.12)** |
| | (p val=0.75, ci [-0.74, 1.64]) | (p val=0.00, ci [0.16, 0.53]) | (p val=0.63, ci [-3.71, 0.54]) |
| School level controls | Yes | Yes | Yes |
| Year fixed effect | Yes | Yes | Yes |
| District fixed effect | Yes | Yes | Yes |
| District-specific linear trend | Yes | Yes | Yes |
| No of Observations | 2,107,532 | 2,107,532 | 2,107,532 |
| R Squared | 0.35 | 0.32 | 0.44 |

Note: The school-level control variables for the second regression include the total number of classrooms, and the availability of electricity and playground in the school. In the first regression, since the dependent variable is total number of classrooms itself, it has been excluded from the list of control variables. The Standard errors are clustered at the district level. For difference in number of observations across Panels A, B and C, see notes to Table 2. ***=significant at 1% level, **=significant at 5% level, *=significant at 10% level. p-value and standard errors using the wild cluster bootstrap method are given in the parenthesis Source: DISE 2009–17; authors' own calculations.

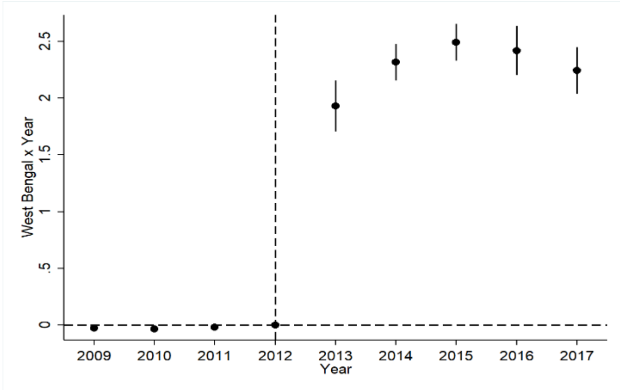
## Appendix

### Figure A1: Event study analysis for pre-primary enrolment

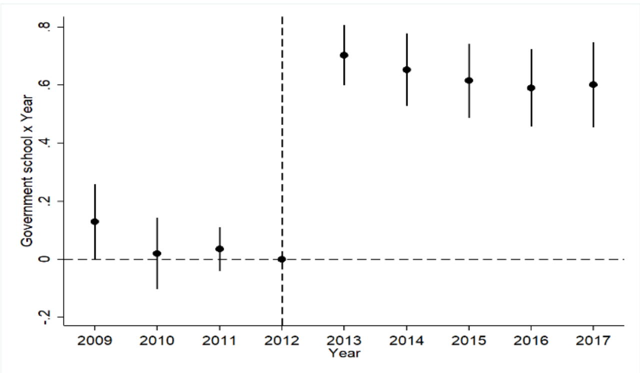
**Panel A: Difference in availability of pre-primary in government schools between West Bengal and control states (DD Specification I)**



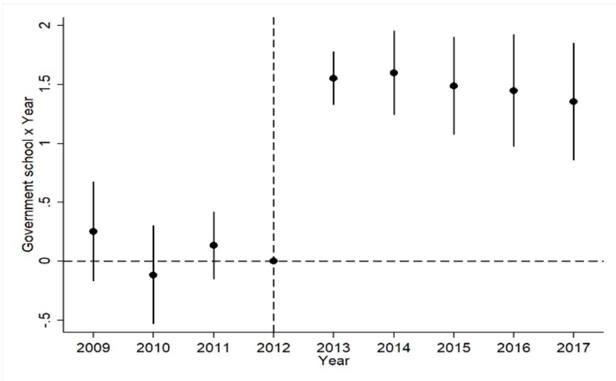

**Panel B: Difference in pre-primary enrolment in government schools between West Bengal and control states (DD Specification I)**



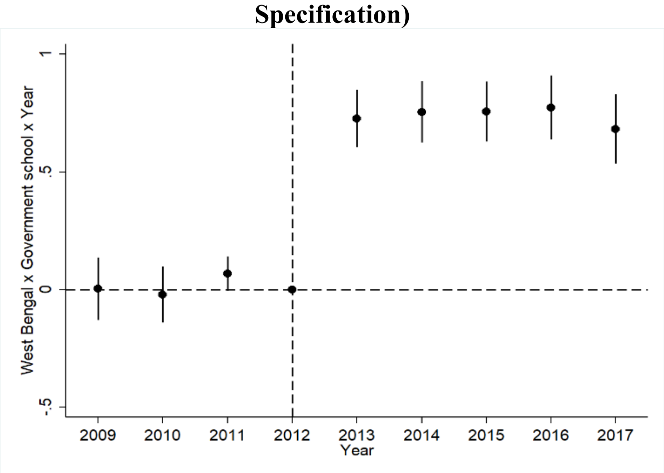

**Panel C: Difference in availability of pre-primary between government and private schools in West Bengal (DD Specification II)**



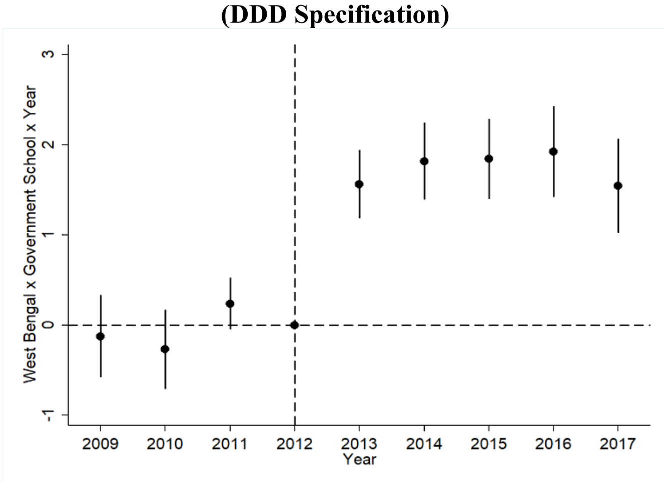

**Panel D: Difference in pre-primary enrolment between government and private schools in West Bengal (DD Specification II)**




**Panel E: Difference of the difference in availability of pre-primary between government and private schools in West Bengal to the same difference in the control states (DDD Specification)**

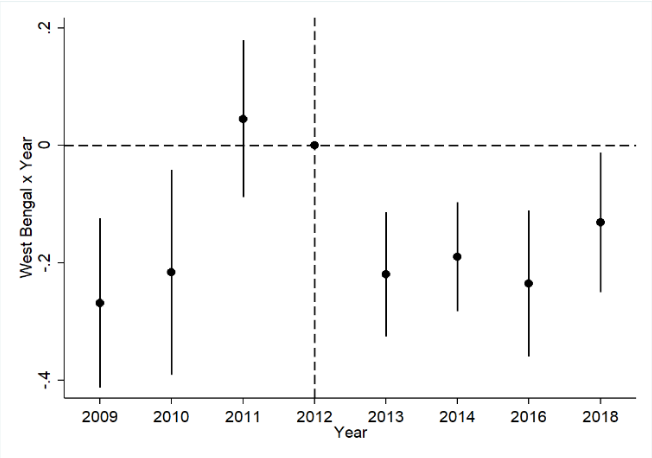



**Panel F: Difference of the difference in pre-primary enrolment between government and private schools in West Bengal to the same difference in the control states (DDD Specification)**

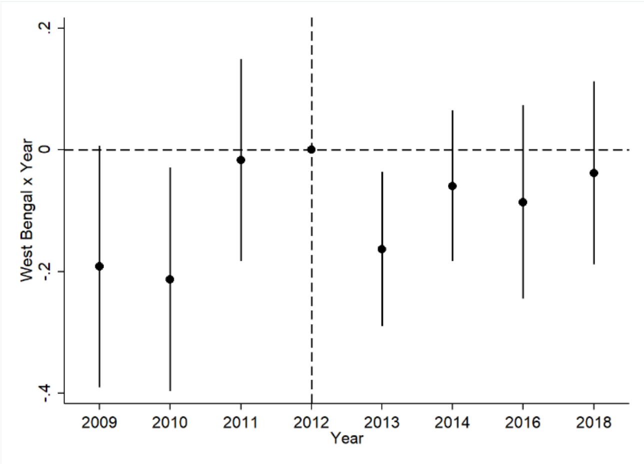


## Figure A2: Event study analysis for learning outcomes of 5–6-year-old children

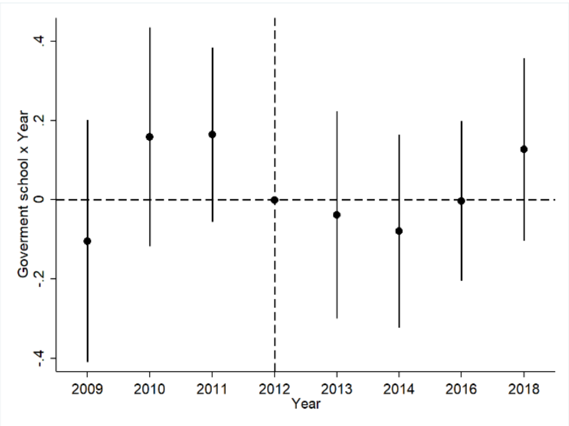
**Panel A: Difference in math score among government school children between West Bengal and control states (DD Specification I)**



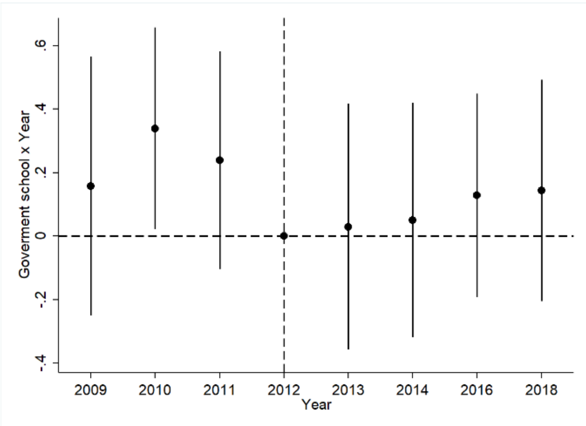

**Panel B: Difference in reading score among government school children between West Bengal and control states (DD Specification I)**



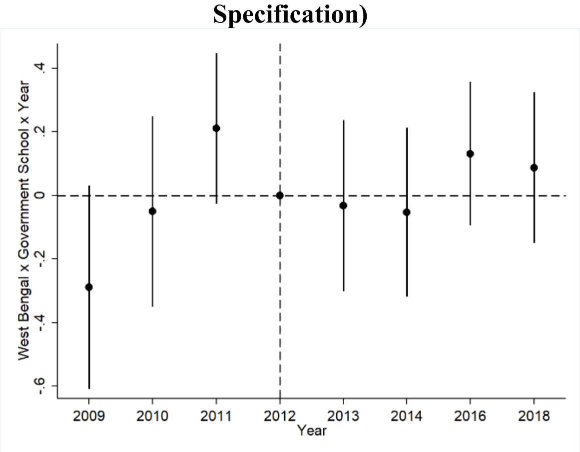

**Panel C: Difference in math score between government and private school children in West Bengal (DD Specification II)**



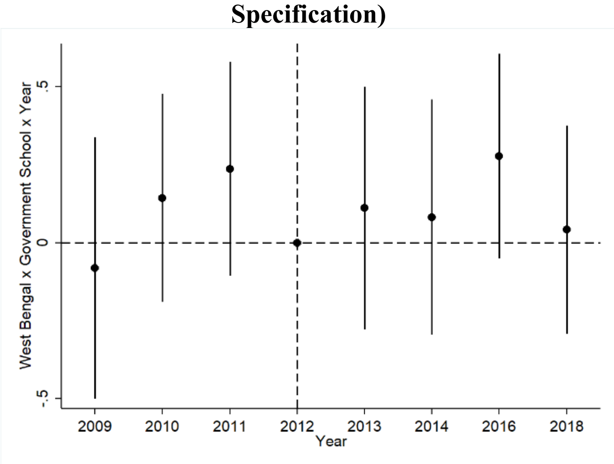

**Panel D: Difference in reading score between government and private school children in West Bengal (DD Specification II)**




**Panel E: Difference of the difference of math score between government and private school children in West Bengal to the same difference in the control states (DDD Specification)**

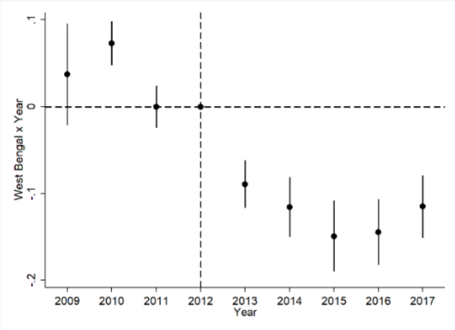



**Panel F: Difference of the difference of reading score between government and private school children in West Bengal to the same difference in the control states (DDD Specification)**

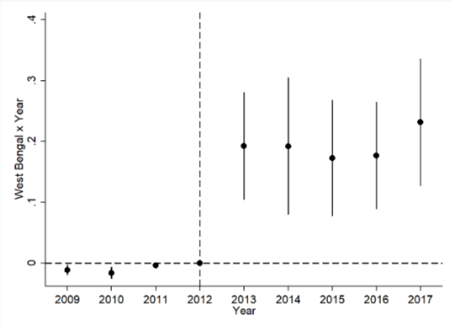


## Figure A3: Event study analysis for school infrastructure

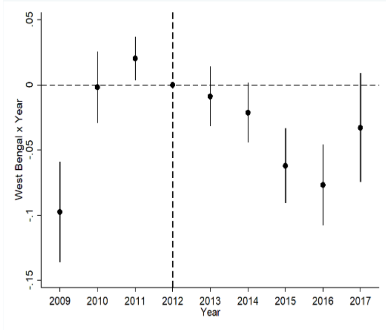
**Panel A: Difference in number of classrooms in government schools between West Bengal and control states**



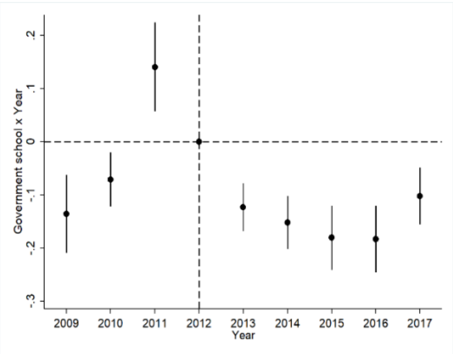

**Panel B: Difference in number of pre-primary teachers in government schools between West Bengal and control states**



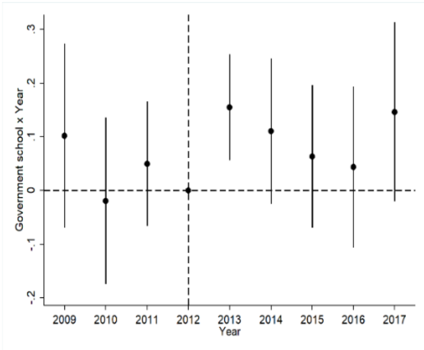

**Panel C: Difference in total teacher in government schools between West Bengal and control states**



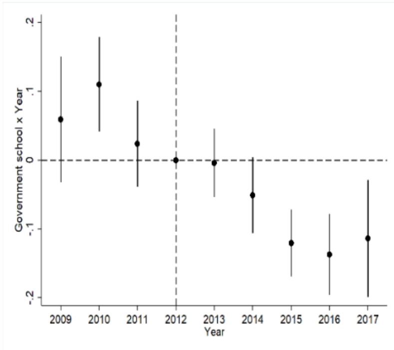

**Panel D: Difference in number of classrooms between government and private schools in West Bengal**



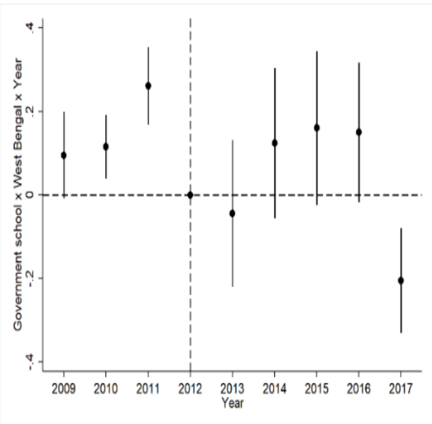

**Panel E: Difference in number of pre-primary teachers between government and private schools in West Bengal**



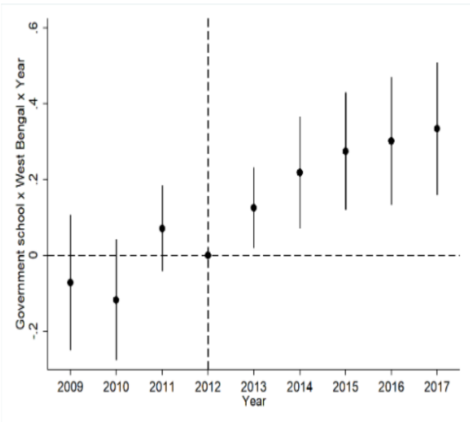

**Panel F: Difference in total teacher between government and private schools in West Bengal**



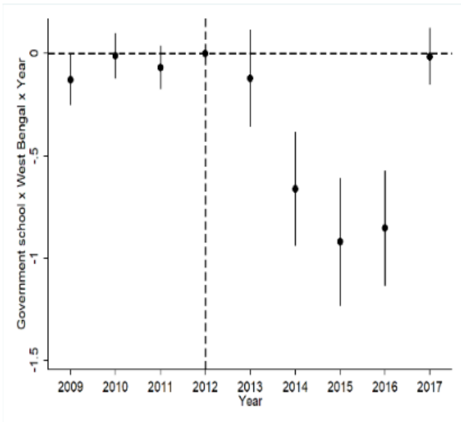

**Panel G: Difference of the difference in classrooms between government and private schools in West Bengal to the same difference in the control states**




**Panel H: Difference of the difference in pre-primary teachers between government and private schools in West Bengal to the same difference in the control states**

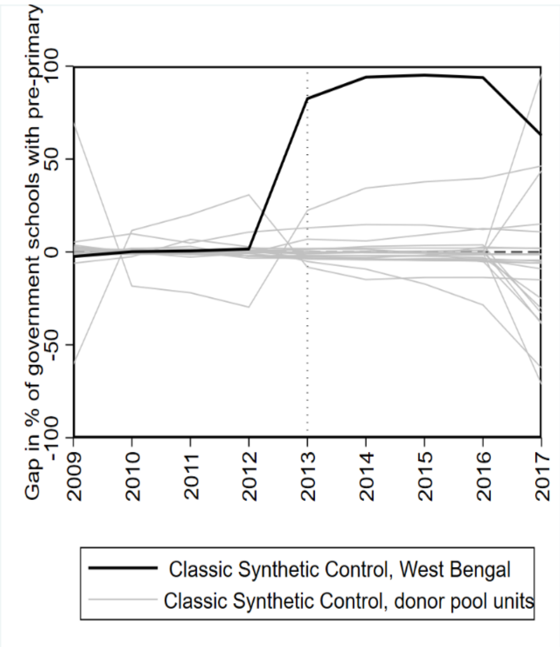



**Panel I: Difference of the difference in total teacher between government and private schools in West Bengal to the same difference in the control states**

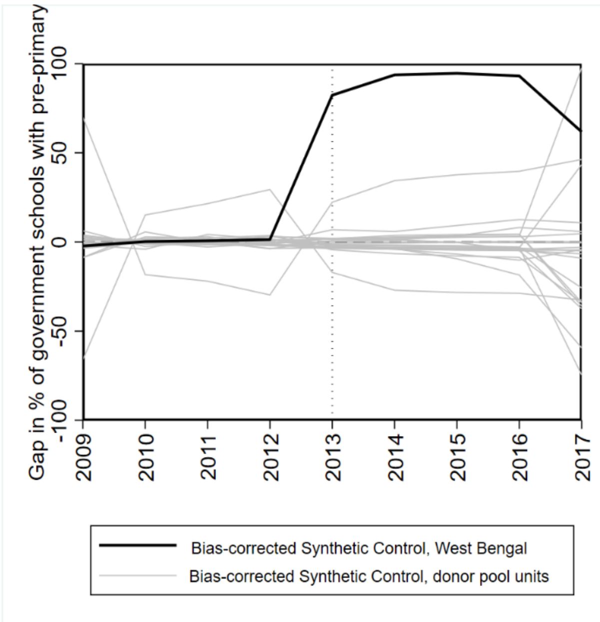


## Figure A4: Synthetic control gap in pre-primary enrolment

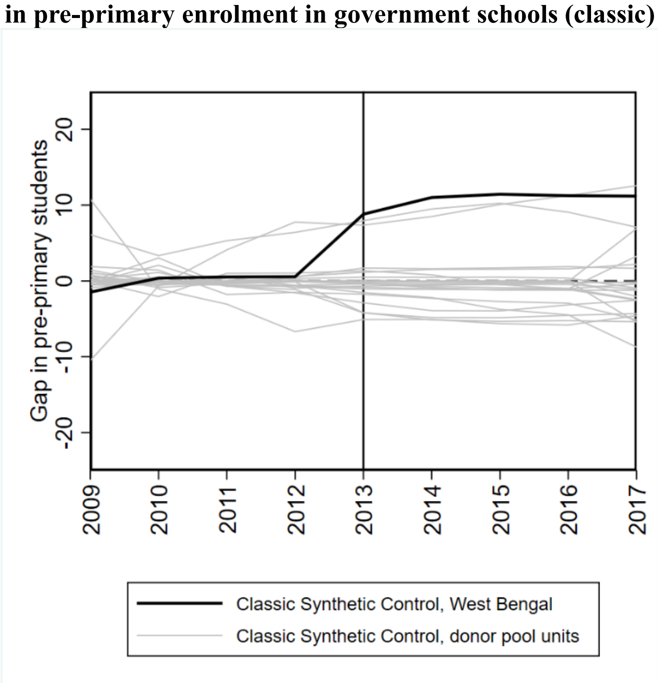
**Panel A: Gap between West Bengal and synthetic control in the availability of pre-primary in government schools (classic)**



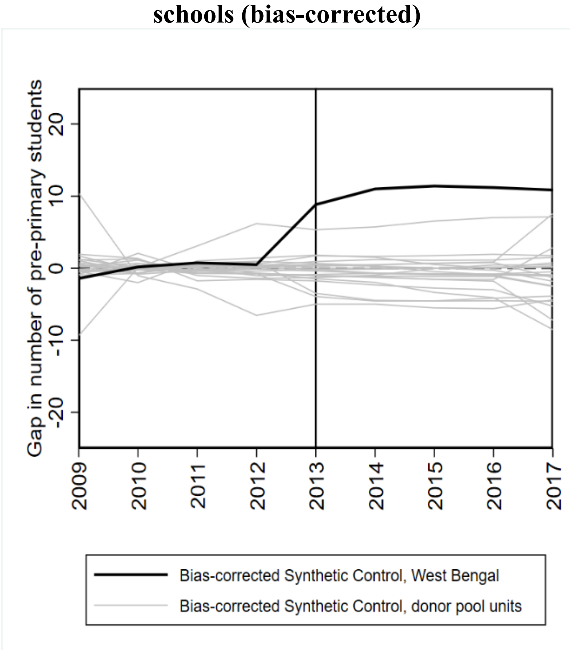

**Panel B: Gap between West Bengal and synthetic control in the availability of pre-primary in government schools (bias-corrected)**




**Panel C: Gap between West Bengal and synthetic control in pre-primary enrolment in government schools (classic)**

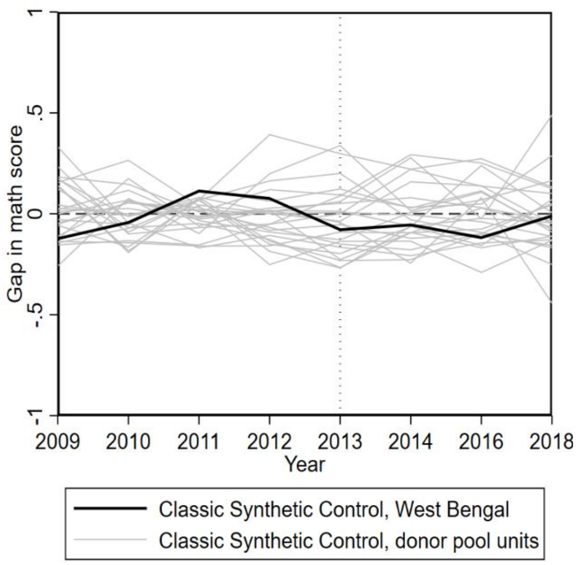



**Panel D: Gap between West Bengal and synthetic control in pre-primary enrolment in government schools (bias-corrected)**

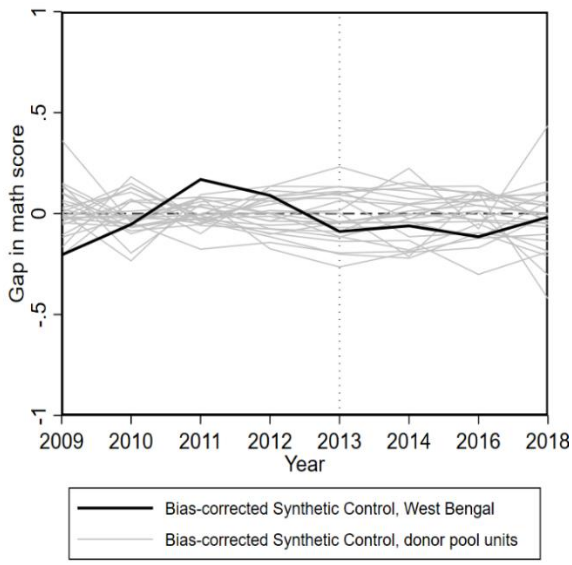


## Figure A5: Synthetic control gap in learning outcomes

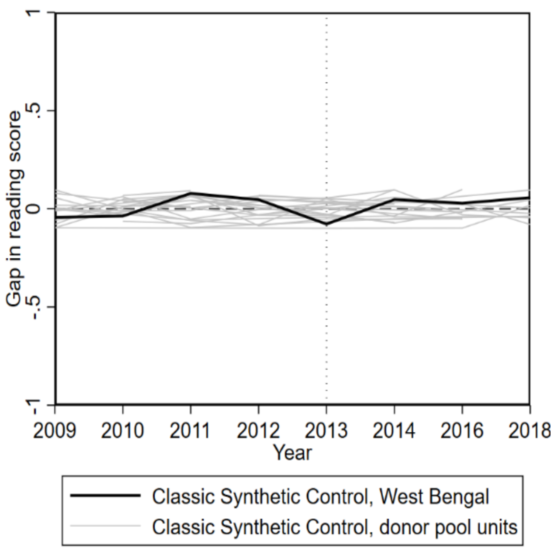
**Panel A: Gap in math score in government school children of West Bengal and synthetic control (classic)**



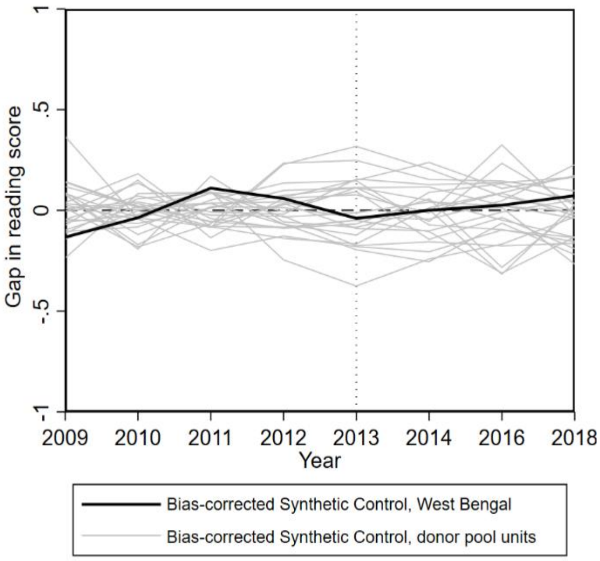

**Panel B: Gap in math score in government school children of West Bengal and synthetic control (bias corrected)**




**Panel C: Gap in reading score in government school children of West Bengal and synthetic control (classic)**

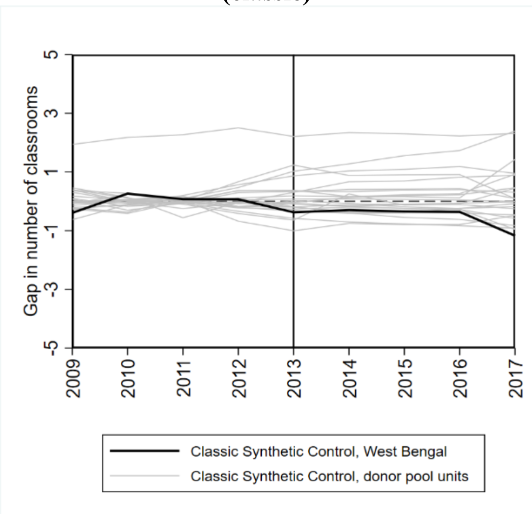



**Panel D: Gap in reading score in government school children of West Bengal and synthetic control (bias corrected)**

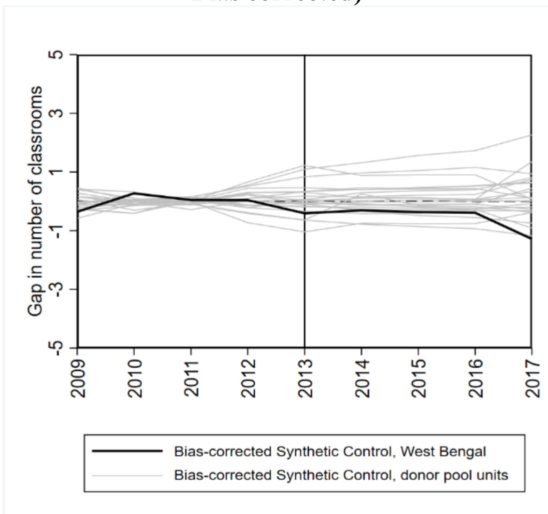


## Figure A6: Synthetic control gap in school infrastructure

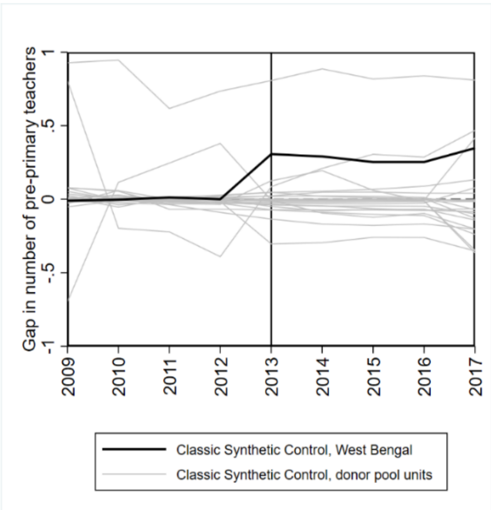
**Panel A: Gap in number of classrooms between government schools in West Bengal and control states (classic)**



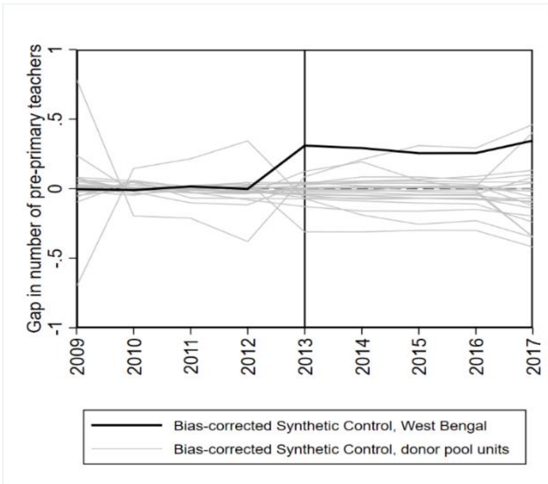

**Panel B: Gap in number of classrooms between government schools in West Bengal and control states Bias corrected)**



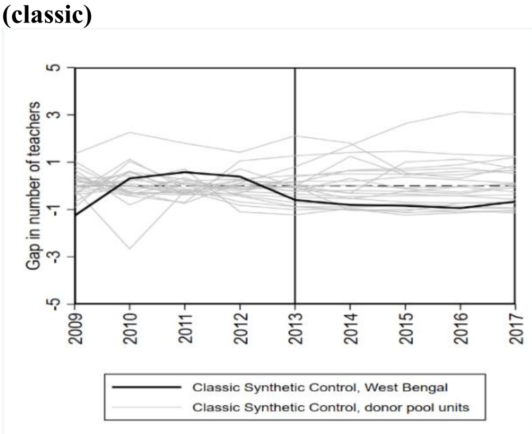

**Panel C: Gap in number of pre-primary teachers between government schools in West Bengal and control states (classic)**




**Panel D: Gap in number of pre-primary teachers between government schools in West Bengal and control states (bias corrected)**




**Panel E: Gap in number of teachers between government schools in West Bengal and control states (classic)**




**Panel F: Gap in number of teachers between government schools in West Bengal and control states (classic)**

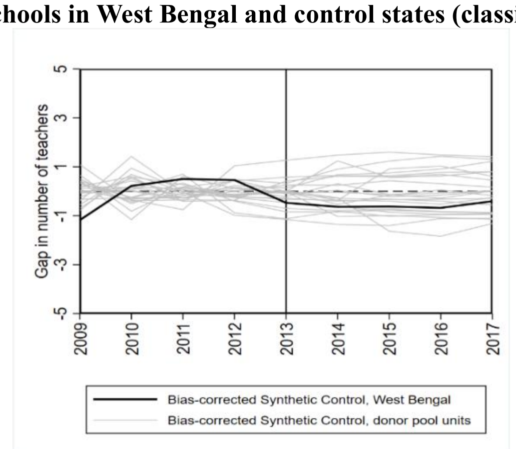


**Table A1: Weights and predictor balance for availability of pre-primary**

Panel A: Weights for the states

| State name | Weight |
|---|---|
| Himachal Pradesh | 0 |
| Punjab | 0 |
| Uttarakhand | 0 |
| Haryana | 0 |
| Rajasthan | 0 |
| Uttar Pradesh | 0 |
| Bihar | 0.17 |
| Sikkim | 0 |
| Nagaland | 0 |
| Manipur | 0 |
| Tripura | 0.47 |
| Meghalaya | 0 |
| Assam | 0 |
| Jharkhand | 0 |
| Orissa | 0.36 |
| Chhattisgarh | 0 |
| Madhya Pradesh | 0 |
| Gujarat | 0 |
| Maharashtra | 0 |
| Karnataka | 0 |
| Kerala | 0 |
| Tamil Nadu | 0 |

Panel B: Predictor balance

| | Treated | Synthetic |
|---|---|---|
| Percentage of schools with pre-primary | 2.20 | 2.24 |
| Percentage of government schools with electricity | 32 | 14 |
| Percentage of government schools with playground | 36 | 43 |
| Average number of pre-primary teachers in government schools | 0.04 | 0.03 |
| Average number of teachers in government schools | 5.55 | 5.54 |
| Average number of classrooms in government schools | 4.34 | 4.31 |

All predictor variables are averaged over the period 2009 to 2012

**Table A2: Weights and predictor balance for number of pre-primary students**

Panel A: Weights for the states

| State name | Weight |
|---|---|
| Himachal Pradesh | 0 |
| Punjab | 0 |
| Uttarakhand | 0 |
| Haryana | 0 |
| Rajasthan | 0 |
| Uttar Pradesh | 0 |
| Bihar | 0.08 |
| Sikkim | 0 |
| Nagaland | 0 |
| Manipur | 0 |
| Tripura | 0.38 |
| Meghalaya | 0 |
| Assam | 0 |
| Jharkhand | 0 |
| Orissa | 0.50 |
| Chhattisgarh | 0 |
| Madhya Pradesh | 0 |
| Gujarat | 0 |
| Maharashtra | 0 |
| Karnataka | 0 |
| Kerala | 0.04 |
| Tamil Nadu | 0 |

Panel B: Predictor balance

| | Treated | Synthetic |
|---|---|---|
| Average number of pre-primary students | 1.31 | 1.30 |
| Percentage of government schools with electricity | 32 | 18 |
| Percentage of government schools with playground | 0.36 | 0.41 |
| Average number of pre-primary teachers in government schools | 0.04 | 0.05 |
| Average number of teachers in government schools | 5.55 | 5.30 |
| Average number of classrooms in government schools | 4.34 | 4.34 |

All predictor variables are averaged over the period 2009 to 2012

**Table A3: Weights and predictor balance for math score**

Panel A: Weights for the states

| State name | Weight |
|---|---|
| Himachal Pradesh | 0 |
| Punjab | 0 |
| Uttarakhand | 0 |
| Haryana | 0 |
| Rajasthan | 0 |
| Uttar Pradesh | 0 |
| Bihar | 0.15 |
| Sikkim | 0 |
| Nagaland | 0.34 |
| Manipur | 0 |
| Mizoram | 0 |
| Tripura | 0 |
| Meghalaya | 0 |
| Assam | 0.38 |
| Jharkhand | 0 |
| Orissa | 0 |
| Chhattisgarh | 0 |
| Madhya Pradesh | 0 |
| Gujarat | 0 |
| Maharashtra | 0 |
| Andhra Pradesh | 0 |
| Karnataka | 0 |
| Kerala | 0.13 |
| Tamil Nadu | 0 |

Panel B: Predictor Balance

| | Treated | Synthetic |
|---|---|---|
| Math score | 1.17 | 1.17 |
| Proportion of pucca households | 0.14 | 0.14 |
| Household size | 6.12 | 6.10 |
| Proportion of households with toilet | 0.39 | 0.55 |
| Proportion of household with electricity | 0.58 | 0.65 |
| Proportion of household with TV | 0.31 | 0.34 |
| Proportion of villages with pucca road | 0.46 | 0.49 |
| Proportion of villages with post office | 0.36 | 0.36 |
| Proportion of villages with bank | 0.19 | 0.19 |
| Proportion of villages with private school | 0.28 | 0.41 |
| Proportion of villages with government primary school | 0.95 | 0.95 |
| Proportion of household where mother went to school | 0.59 | 0.68 |

All predictor variables are averaged over the period 2009 to 2012

**Table A4: Weights and predictor balance for reading score**

Panel A: Weights for the states

| State name | Weight |
|---|---|
| Himachal Pradesh | 0.37 |
| Punjab | 0 |
| Uttarakhand | 0 |
| Haryana | 0 |
| Rajasthan | 0 |
| Uttar Pradesh | 0 |
| Bihar | 0 |
| Sikkim | 0 |
| Nagaland | 0.21 |
| Manipur | 0.06 |
| Mizoram | 0 |
| Tripura | 0 |
| Meghalaya | 0 |
| Assam | 0.06 |
| Jharkhand | 0.24 |
| Orissa | 0 |
| Chhattisgarh | 0 |
| Madhya Pradesh | 0 |
| Gujarat | 0 |
| Maharashtra | 0 |
| Andhra Pradesh | 0 |
| Karnataka | 0 |
| Kerala | 0.06 |
| Tamil Nadu | 0 |

Panel B: Predictor Balance

| | Treated | Synthetic |
|---|---|---|
| Reading score | 1.19 | 1.19 |
| Proportion of pucca households | 0.14 | 0.23 |
| Household size | 6.12 | 6.17 |
| Proportion of households with toilet | 0.39 | 0.56 |
| Proportion of household with electricity | 0.58 | 0.82 |
| Proportion of household with TV | 0.31 | 0.51 |
| Proportion of villages with pucca road | 0.46 | 0.49 |
| Proportion of villages with post office | 0.36 | 0.35 |
| Proportion of villages with bank | 0.19 | 0.17 |
| Proportion of villages with private school | 0.28 | 0.33 |
| Proportion of villages with government primary school | 0.95 | 0.90 |
| Proportion of household where mother went to school | 0.59 | 0.68 |

All predictor variables are averaged over the period 2009 to 2012