Knowledge is Power: School Construction & Intergenerational Human Capital

Naveen Sunder*

Economics Department, Cornell University

December 16, 2017

Abstract

This paper examines how a school building program in India impacted educational and cognitive outcomes of the children of program beneficiaries. Beginning in the early 1990s, India implemented a national school building programme which disproportionately benefited districts that had low female literacy levels at the time. Records indicate that the programme covered nearly 375,000 schools and benefited more than 50 million children across 271 programme districts. Using a regression discontinuity approach, I exploit spatial and temporal variation to estimate the intergenerational effects of improvements in schooling opportunities experienced by women under this scheme. For the analysis, I combine data from several sources: I use individual-level data from multiple nationally representative datasets, government archival data on programme implementation at the district-level, and detailed administrative data on school infrastructure. I find that children whose mothers were impacted by this programme had higher cognitive abilities as measured by scores on standardized math, reading and English tests. These children also experienced positive impacts on enrolment and grade progression. While I observe that these educational benefits exist for children of both genders, the impacts on test score for girls is between 13-32 percentage points higher than the corresponding effect for boys. I explore different mechanisms through which school construction might have led to the child-level effects I observe, such as maternal education, health and healthcare utilization. The findings of this paper provide evidence on how the expansion of school infrastructure could have substantial long term impacts on cognitive and educational outcomes that extend beyond the generation that is directly impacted by the school construction.

JEL classification: I21, I25, O15, J24

Keywords: Education, Economic Development, Human Capital, India, Regression Discontinuity

*Address: B16 MVR Hall, Cornell University, Ithaca, New York- 14853 , e-mail: *fn63@cornell.edu*. I thank Prateek Bansal, Averi Chakrabarti, Heidi Kalia, George Jakubson, Prabhu Pingali and David Sahn for thoughtful discussions and useful comments on earlier versions of this paper. All remaining errors are my own. **Preliminary draft - Please do not cite or distribute without prior permission.** The earliest draft of this paper was dated: 4 April 2016.

1 Introduction

The recent expansion in educational opportunities across the globe has been impressive. Educational inputs (like physical infrastructure) have rapidly increased and so have outcomes such as net enrollment rates. Compared to the formative stages of the currently industrialized nations, the current pace of increase in different educational outcomes in developing countries has been far more rapid. According to estimates in Barro and Lee [2013], the share of population without any formal schooling has fallen from 54.6 percent in 1960 to around 17.4 percent in 2010. This is a significant achievement considering the population of these countries has nearly tripled during this period, from 1.36 billion to 3.95 billion. A large part of this change has been driven by progress in Sub-Saharan Africa (72.3 percent to 32.2 percent) and South Asia (73.5 percent to 33.3 percent).

India has managed great strides in the education sector since the early 1990s. The total literacy rate increased from around 43 percent in 1981 to nearly 74 percent in 2011 (Office of the Registrar General & Census Commissioner, India). Although these improvements were not equally experienced by both genders, the gender gap in literacy has declined over this time period. The Gross Enrollment Ratio (GER) in India has also seen a steady upward trend, rising from a a relatively low 42.6 percent in 1951 to 115.5 percent by the year 2011. This growth in enrolment and literacy has been supported and spurred by the rise in the number of educational institutions at different levels of education. From having close to 330,000 primary education facilities in the country, India now has close to 847,000 primary education facilities[1]. A similar rise has been experienced in infrastructural facilities at higher levels of education - In the period between 1981 to 2011, the following trends have been recorded- Upper Primary (118,600 to 478,800) Secondary (51,600 to 84,100), College (696,300 to 3,297,400) and Universities (11,000 to 62,100).

The improvement in physical infrastructure in Indian education has not, however, been accompanied by an equivalent enhancement in the quality of education and the amount of learning taking place in schools. *Pratham* is a not-for-profit organization tbat administers standardized tests at a national level to assess the reading and math skills of children in grades 1 to 8. According to a national survey conducted by them in 2016, even by the time the students reach grades four or five, they are not able to answer questions that are meant to be answered by students in second grade. According to the same data (ASER 2016 report), only 47.8 percent of children in

[1] Data accessed on 09 October 2017 at link

grade five could read at a grade two level. The corresponding number for students in grades two and three are 13.4 and 25.1 percent respectively. This implies that by second grade, nearly 87 percent of children are already behind on the reading skills that they should have achieved. Similar patterns are observed in arithmetic skills as well, where nearly 18 percent of students in second grade could not even recognize one digit numbers. The same problems get carried over to higher grades- nearly 50 percent of all children in grade five could not solve a simple subtraction problem.

XXX Potentially discuss the importance of childhood cognitive skills over here.

Early academic performance is critical for the long term human capital development of children. The development of cognitive abilities depends on the school environment and quality along with home and parental inputs, among other factors. Using data from Senegal, Glick and Sahn [2010] find that second grade test scores (which are expected to proxy for cognitive skills) have an impact on test performance seven years later. They also provide evidence that differences in test scores stemming from early human capital deficiencies might lead to higher grade repetition, which in some cases might cause children to drop out from school. Kaila et al. [2017] use panel data spanning 15 years from Senegal and Madagascar to find that second grade test scores have a significant impact on adult cognitive skills as measured by French and math test scores. These effects remain even after controlling for assets, health and school-level characteristics of children during second grade. These analyses demonstrate the influence that early academic performance has on later life outcomes, highlighting the importance of learning at a young age and the relevance of quality education in primary schools.

This problem is further compounded by the large inter-state differences within the country. While in Himachal Pradesh (a state in India) nearly 57 percent of children in grade three were able to solve a subtraction problem, while the corresponding number in Madhya Pradesh (another state) is close to 14 percent. These kinds of disparities indicate that the quality of learning (as measured by these tests) has not shown the considerable progress made in terms of physical infrastructure and enrolment outcomes. This sort of a learning shortfall would exacerbate the already existing inequalities that are present between these regions. This in turn, severely affects the labor market (and other) opportunities that youth have access to (as discussed earlier).

How can mother's education impact different child level outcomes? Write about that literature and then tie it into the impact of women's education on child education quality. Would be a good segue into the next paragraph

In this paper, I estimate how enhanced educational opportunities experienced by women in their childhood have an effect on their children's human capital development. A national school building programme that took place in India in the mid to late 1990s provides the exogenous variation in school exposure that I use for this analysis. India's District Primary Education Programme (DPEP) was a school scheme targeted towards districts with relatively poor educational outcomes. Under this programme, school building was targeted towards districts where female literacy was below the national average female literacy rate at the time (39.2 percent). If the assignment rule had been followed perfectly, the setting would have provided a sharp discontinuity in the probability of treatment around the cutoff (the national average female literacy rate). However, when the programme was actually implemented, the rule was not followed perfectly (discussed later), but well enough to provide a Fuzzy Regression Discontinuity (FRD) empirical setup.

I measure children's human capital by focusing on outcomes related to child-level schooling related outcomes. Typically, schooling outcomes are measured in terms of enrolment or number of years of education attained. While I do measure impacts on these variables, I also look at the quality of human capital by focusing on the realized learning outcomes of the children. I do this by using test scores on English, reading and math tests from a large nationally representative sample of students who were administered this test. In measuring human capital using these outcomes, I examine what these children have learned over the years in school and elsewhere, and not merely outcomes like whether children are going to school (or not) or how many years of schooling they have completed.

I find improved educational outcomes for the children of women who went to school around the time the school construction programme was implemented. These children have higher cognitive abilities as measured by scores on reading, math and English tests. On the extensive margin, they were more likely to answer at least one of the questions on these tests correctly. In addition, I find positive effects of the programme on child-level school enrollment and grade progression outcomes. These children also had lower rates of dropout as compared to the children in the control group. Having estimated the intergenerational effects of the school construction programme, I explore potential mediators of this effect. I find evidence that the school construction programme had a significant impact on the education level of the women who were of school going age in the districts that benefited from the programme. This kind of schooling boost could have potential knock-on effects on other socio-economic outcomes for these women. In the current setting, I find a positive impact on women's health (Grossman [2015]; Grépin and Bharadwaj [2015]; Agüero and Bharadwaj [2014]; Amin et al. [2013]), contraceptive usage (Johnston et al. [2015]; Andalón et al. [2014]) and usage of facilities related to Ante Natal Care (ANC) and Post Natal Care (PNC). I discuss the potential role that these factors could have played in mediating the effect of DPEP on intergenerational human capital.

The remainder of this paper is organized in the following manner: I start with a review of the literature (section 2) and in section 3 I discuss the DPEP programme in detail. Section 4 lays out the empirical strategy and explains the classification of the sample into treatment and control groups, while section 5 focuses on the data used in the analysis. In section 6, I present my main results along with the robustness and falsification checks. I finish with a discussion on potential mechanisms (section 7) and conclude in section 8.

2 Literature Review

This paper adds to the limited literature that estimates the impact of expansion of school infrastructure on different outcomes. In one of the early papers, Duflo [2004] explores the impact of a school construction program in Indonesia in the 1970s on education and wages and finds that years of education went up by 0.12 to 0.19 years for an increase in density of schools (by one school per 1000 children). In addition, she finds that the rate of return on education is in the 6-10 percent range, depending on the specification that she uses. Using a randomized controlled trial in 31 villages in Afghanistan, Burde and Linden [2013] find that increased access to schools at the village level has positive impacts on enrolment and test scores. They find that the programme has a larger impact on girls than on boys, which almost leads to the elimination of the previously existing gender differences in these outcomes.

Another school construction programme whose impacts have been causally evaluated is the BRIGHT school construction in Burkina Faso. This programme specifically constructed "girl friendly" schools in rural areas of Burkina Faso. The programme was allocated based on an index that was created for 293 villages, with villages in the bottom half of the distribution receiving the programme. Kazianga et al. [2013] find that the school construction had a positive impact on both enrollment and test scores of programme beneficiaries. They find that the unique characteristics of the new schools that had been constructed played a critical role in this effect, but find that merely the presence of the schools also had a statistically significant impact on these outcomes. Using a similar econometric setup, De Hoop and Rosati [2014] find that although the programme led to higher rates of enrolment, it did not have a significant effect on child labor in these regions. In fact, they find that for some sub-groups child labor might have increased as a result of the programme.

The impact of school construction programmes has also been probed in several developed country settings. Cellini et al. [2010] estimate the impact of such policies on housing prices, school performance and district composition. Using a public referendum in the US on the issuing of bonds to facilitate enhanced public school expenditure, the authors find that marginal home buyers are willing to pay more in districts with higher public spending. These increases in price cannot be full explained by commensurate increases in the quality of educational outcomes as measured by test scores. Neilson and Zimmerman [2014] study the impact of school construction in a poor urban school district in the US and detect positive impacts on child enrolment and test scores, as well as on home prices.

There so exist some papers that have evaluated the impact of the DPEP programme that I focus on, but they discuss different socioeconomic outcomes. Khanna [2015] estimates the general equilibrium effects of this school expansion on the returns to education. He finds that the partial equilibrium effects are larger than the general equilibrium ones and that there are differential effects by cohort and skill level. Azam and Saing [2016] find that the programme leads to higher enrollment, number of years of education and probability of completing primary education among those directly impacted by the programme. In an earlier assessment of the programme, Jalan and Glinskaya [2013] find small effects on enrolment only for socially disadvantaged groups, with very little narrowing of the gender gap. However, this paper only looked at the impact of the programme on the 42 districts that were a part of the first phase of the project, and not on all the districts that were eventually part of the school construction treatment.

The papers discussed above have some parallels with the current analysis, but there are critical differences that distinguish this work. A key innovation of my analysis is that I estimate the intergenerational impact of the school construction scheme, rather than the impact of the DPEP programme on the population directly exposed to it. In my knowledge, there are no other papers that have explored this aspect of the impact of school construction policies. Since these programmes create infrastructure that remains in place for a long period of time, their costs need to be assessed viz-a-viz their long term impact on different cohorts, and not only their direct impact on the beneficiaries. By providing evidence on the impact of such policies on the children of the beneficiaries, I provide arguably the first piece of evidence that quantifies the intergenerational benefits of school construction.

Secondly, I use information provided by government archives to define more precisely the timing of initiation of the programme. The DPEP programme was implemented throughout India in broadly three phases. The papers discussed above either use a uniform start year for the programme across all districts in a Regression Discontinuity (RD) framework (Khanna [2015]) or use information on the different phases of the programme in a Differences-in-Differences (DID) approach (Azam and Saing [2016]). Even among districts within a certain phase of the programme, there were variations in the exact year in which a district received the scheme. I obtain and use the exact start year of the programme in each treatment district (around 271 in number) by manually analyzing paper archival records of the government of India. This process assists me in defining the treatment/control groups more accurately.

Further, in the paper I explore in detail the potential mechanisms through which the observed impacts might be mediated. For this purpose, I combine information from different nationally representative household surveys in a way that has not been done before in the literature. This allows me to look at a variety of education and health related mechanisms that might have played an important role in mediating this effect.

3 DPEP Programme

The Indian Constitution, which was adopted in 1950, classified education as a state subject, which made states responsible for the majority of funding going into the education sector. However, in 1976 a constitutional amendment placed education under the concurrent list, enabling both the central and the state governments to create legislations, provide finance and create policies in the educational sphere. In 1986, the National Policy on Education (NPE) "resolved to achieve the goal of Universal Elementary Education (UEE) by the year 2000". In a continued push towards education reform, the governments at the state and the central level implemented a host of policies in the 1980s. The most important such policies were - *Andhra Pradesh Primary Education Project*, *Shiksha Karmi Project*, *Mahila Samakhya Program*, *Bihar Education Project*, *Lok Jumbish Parishad Program*, *Operation Blackboard*, and the *Uttar Pradesh Basic Education Project*. The District Primary Education Programme (DPEP) evolved from these projects and was designed keeping in mind the lessons from these other policy efforts.

The DPEP school building programme was implemented across India in a phased manner. It introduced a decentralized design and flexibility, while refocusing efforts towards holistic approaches, a focus on equity, quality and capacity building in Indian education (Pandey [2000]). The first phase was launched in 42 districts in 1993. Eventually, the programme was implemented in 271 districts across 18 states in India. With an aim of improving education and literacy, DPEP was allocated to districts that were relatively backward on certain educational outcomes. The allocation rule for the programme was based on two criteria. The programme was to be made available to those districts in which the female literacy rate (DFLR) was less than the national average as measured by the 1991 census. The 1991 figure of 39.2 percent, based on the most recently available census data, was chosen as the benchmark as it was the most recently available nationally representative education data at the time of programme implementation. The government also wished to introduce the programme in districts only if they had successfully implemented the total literacy campaign, but this criterion did not end up being a basis of selection since by 1994, the TLC had been implemented successfully across all districts in India (Jalan and Glinskaya [2013]).

The cost of the DPEP policy was shared between the central and the state government in a 85:15 ratio. The central government of India received support for this programme from various donor organizations like the Official Development Assistance (ODA), the Royal Government of the Netherlands, the U.K. Department for International Development (DFID) and the World Bank. In order to ensure that the new programme did not crowd out funds that were already being spent on education related policies, the central government stipulated that the states had to contribute their share of the DPEP funding while maintaining their existing level of education expenditures. Since the DPEP funds were committed over and above the already existing education budget, the programme represented a big surge in education expenditure across the country.

DPEP was very effective in improving the education infrastructure. By the year 2000, the programme had led to the opening of more than 1,60,000 new schools and to the creation of nearly 84,000 alternative schooling (AS) centres (Azam and Saing [2016]). Over a decade, these AS centres had catered to more than three million children. The programme also led to large scale hiring of teachers. Estimates show that about 1,77,000 teachers were hired and nearly one million teachers were trained under this programme. According to Jalan and Glinskaya [2013], the programme potentially impacted close to 51.3 million children across 375,000 schools in the country. In addition, given the decentralized nature of DPEP implementation, the programme increased the local presence of monitoring and implementation agencies - more than 3,000 block resource centers and nearly 30,000 cluster resource centers were established under the DPEP scheme.

Pandey [2000] identifies several drivers that led to the relative success of the DPEP programme, such as a strong focus on student learning, decentralization and local empowerment, and constant attention on building capacity. These were important changes as they represented a shift in the nature of education policy in the country. This was also the first time in India, the district was made the unit of planning for education instead of the state. Such shifts facilitated the successful implementation of this programme across the country.

4 Empirical Strategy

4.1 Identifying Treatment Units

After applying several criteria, I divide the sample into treatment and control groups. I discuss below each of these criteria in further detail.

Women's district of schooling: Ideally, the treatment status of a woman (direct beneficiary of DPEP) would be determined on the basis of the district in which she resided when she was of school going age. Unfortunately, the datasets I use do not provide this information. Therefore, I use the current district of residence to assign treatment/control status to each woman (mother) in my sample. The potential issue with this assignment mechanism is that the current district of residence may not be the same as the district a woman lived in when she went to school. This is of particular concern in a context like India since women tend to migrate to the husband's house after marriage. Therefore, using the current district of residence could potentially be an issue if there was extensive post-marriage inter-district migration amongst women. Evidence from India, however, shows that inter-district migration is generally low. Using National Sample Survey (NSS) data from 1983, 1987 and 1999, Topalova [2007] estimates that overall inter-district migration is low - only three to four percent of people report any migration across district lines in India. Additionally, Topalova [2007] demonstrates that although nearly 40 percent of rural women report a change in their location after marriage, a very small proportion of this movement is across district boundaries.

This is further supported by the findings of Bloch et al. [2004], who in their paper report that in India the average distance that a woman moves after marriage is about 21 miles. Given that in 2001 the average size of a district in India was close to 2100 squared miles, most post-marriage migration was likely to have taken place within districts rather than across districts. Combining these findings, it could be reasonably expected that women who married in and around the decade after 1990 would most probably have resided in the same district in which they spent their schooling years. Hence assigning treatment status based on current district of residence should not lead to significant misclassification errors in treatment status.

Women's school going age: The sample used in the analysis is based on women who were of school going age in the treatment and control districts at the time of DPEP programme implementation. Evidence suggests that in India in the 1990's there was a sharp decline in school enrollment at the age of 18 years (Khanna [2015]). Hence, this is a candidate for a cutoff age to define the women who could have been impacted by this programme[2]. Since the focus of DPEP was mostly on increasing access to primary schools, for my main analysis I use a lower cutoff of 14 years to identify women who were of school going age at the time of programme implementation. This 14 year age cutoff is also consistent with the fact that although the ideal primary school going age is 6-10 years, children up to the age of almost 13 years stay in school due to delayed school entry or uneven grade progression (Azam and Saing [2016]).

There is evidence that girls tend to drop out of schools at slightly younger ages than boys. This may be due to a variety of reasons, chief among them being child marriage and onset of menarche. The effect of onset of menses on schooling attainment has been studied in developed (Burrows 1 and Johnson [2005]; Roberts et al. [2002]; Joan and Zittel [1998]) and developing coun-

[2] I use this cutoff as a robustness check of my results

tries (Sommer [2010]). Other evidence from India finds that girls missed a class or had to refrain from other kinds of work due to the onset of the menstrual cycle (Sharma et al. [2008]). The same study finds that menstruation affected the daily routine of nearly 60 percent of the girls in their sample. Therefore it seems likely that menstruation further exacerbates the dropout rates for girls, especially in the absence of developed sanitation facilities in school. Therefore, it may be the case that onset of menarche might have led to higher and earlier dropout from schools amongst girls (Adukia [2017]; Kirk and Sommer [2006]; Burgers [2000]; Fentiman et al. [1999]). Therefore, in order to assuage any concerns related to a higher age cutoff (14 years) driving my results, I conduct robustness checks where I consider a lower cutoff (12 years) to define the group that is at risk of schooling *risk of schooling*.

Timing of the programme-Treatment Districts: Apart from the age of individuals, the temporal aspect of implementation also restricted participation in the programme. DPEP started being implemented in the year 1993-94. One way to define the sample for analysis would be to include all women of *school going age* in the treatment and control districts in 1993-94. This would imply that each district would be assigned the same start year for the programme, which is the approach taken by Khanna [2015]. I potentially improve on this specification by using the *exact* timing of DPEP implementation across different districts. School construction programmes by their very nature are large scale infrastructure projects and take a considerable amount of time and expertise to be implemented. Given these challenges inherent in this task, the timing of DPEP implementation varied dramatically across different districts. The school construction was carried out in phases and took a few years to get started in some districts. Using archived government documents, I estimate the *exact* year of initiation of DPEP in each treatment district and use it in my empirical strategy.[3]. Instead of using a standard start date of the programme, I use this *exact* start date in my empirical strategy.

In figure 6, I provide examples to illustrate the accuracy gained due to the information I infer from the government archives. Since the DPEP launched a large scale construction of schools in the treatment districts, one would expect a rise in the rate of growth of schools in or around the time that the programme was implemented. In the figure, I use detailed school level information from the DISE data to plot the year on year rate of growth of schools over time in several districts.

[3]I manually went through thousands of pages of government documents to assign start dates of the programme for each treatment district

The red line (dashed) represents the year when the DPEP programme was meant to begin (1993-94) in all the treatment districts and the black line (solid) shows the year in which the programme appeared to be implemented in a particular district according to the archived government records.

The graphs illustrate that the spike in the rate of growth of schools in a particular district is better predicted by (or is *closer* to) the start year determined from the government documents (solid black line), rather than the uniform start year of 1993-94 (dashed red line). Recall that the source of the information for the rate of growth of schools (DISE data) is different from the archival documents used to infer the start dates in the treatment districts, and that the data come from different points in time - DISE data from 2005 and archival documents from 1992-2000. Despite these differences, the two sources still seem to present a unified story about the implementation of the DPEP programme. This further adds credibility to the information inferred from the archival documents while clearly indicating that the data from the government archives is closer to the *actual* start date of the programme in the treatment districts. The use of the archival data plausibly helps in defining the treatment/control districts in a more *precise* manner. This suggests a significant value addition in the usage of this archival data in assigning separate starting years for each treatment district.

Timing of the programme-Control Districts: By definition, the control districts do not receive the programme and hence there is no obvious start year of the programme in these districts. Still, there is a need to assign a starting year in these districts as I have to choose the correct comparison population within these districts. I assign a starting year of the programme for these districts using two different methods. Under the first method, I take the average of the starting year of the programme in the treatment districts within a particular state and assign this date as the start date for all control districts within the state. As the aim of the analysis is to find the population in the control group that would be of school going age around the same time as the programme was being implemented (in treatment districts), using a more *local* definition of timing of implementation seems logical. therefore, I use this definition of start date assignment to control districts in my main specifications. Another way to impute the starting year would be to assign all control districts the nationwide average start date among the treatment districts (instead of using the state averages). This method ignores the state (or regional) differences in implementation patterns across different districts. To show the robustness of my results, I replicate the main results using this alternative definition to identify the control group.

Sample of Children: After identifying the women who would compose the treatment/control group, I turn my attention to the identification of the children who will be part of the estimation sample. Here it is important to keep in mind that the main mechanism through which I argue that the programme has an intergenerational effect is through the differences in schooling opportunities experienced by the mothers of these children. Hence, I need to identify a group of children who themselves were not *directly* impacted by the programme, but rather experienced the impact of the programme only through the effect it had on their mothers. I use the fact that the DPEP programme ended around the year 2001-04[4] to identify a cohort of children who would be of school going age only after the DPEP programme had come to a stop.

I restrict the sample of children in my analysis to those kids who would have potentially began their schooling in the year 2005 or beyond. Since schooling typically begins at the age of 5 or 6 years, I restrict the sample of children in my analysis to those who were born in the year 2000 or afterwards. Figure 10 depicts the distribution of the year of birth of the children in the main estimation sample. Figure 11 graphs the same information separately for the treatment and control group. As the figure indicates, there are no noticeable differences in the age distribution of the children in the treatment and control groups. This alleviates concerns of children's age and other cohort related effects driving the main results.

Changes in District Boundaries: The DPEP programme was implemented at the district level[5]. As discussed earlier, the number of districts in India has been growing over the past four or five decades. From around 356 districts in the 1971 Census, there were approximately 640 districts as per the 2011 Census. This implies that the number of districts as per the 1991 census (466 districts) is considerably different from the number of districts in the 2001 (592 districts) and 2011 (640 districts) Census surveys. KUMAR and SOMANATHAN [2009] provide a detailed discussion of the different types of changes in district boundaries that happen during this period. I use their district concordance tables assist in mapping districts across the different waves of the census. I extend their analysis to match districts that were created after the 2001 Census to the districts that they correspond to in the 1991 Census.

[4]This was the time that the national level Sarva Shiksha Abhiyan (SSA) was introduced. This was introduced in all districts of the country and aimed at expanding education opportunities across the whole nations. In that respect it was different from the DPEP which targetted the programme based on the allocation rule.

[5] This level is akin to counties in USA. In India, the district level is one step below the state level in the administrative heirarchy.

There were different kinds of changes in district boundaries across these census years. In cases where new districts were created from a single parent district, I assign the treatment status of the parent district to districts carved out of them. There are some cases in which a district was formed from two or more parent districts. If *all* the parent districts have the same treatment status, then the newly created district is also assigned the same treatment status. But, it might be the case that the treatment status of the *parent* districts differ, in which case I assign the treatment status to the new district based on whether a major share of its population comes from the control or the treatment parent district. In other words, if more than 50 percent of the population of the new district comes from the parent districts of a certain treatment status, I assign this treatment status to the new district. I use analogous rules in assigning programme start years and district female literacy rates (1991 Census) to the newly created districts.

4.2 Methodology

Having assigned the treatment status, one could estimate the relationship between the DPEP programme and an outcome such as child test scores by using an Ordinary Least Squares (OLS) approach. This would take the following form :

$$Y_{idt} = \beta DPEP_d + \delta X_{idt} + \epsilon_{idt} \tag{1}$$

where, Y_{idt} represents the outcome of interest at the individual level, $DPEP_d$ represents whether the individual belongs to the treatment or control group and X_{idt} represents a vector of individual level control variables used in the regression. Here β would be the coefficient of interest. But, β, the coefficient on the $DPEP_d$ variable, may not provide the *causal* impact of the programme, for example due to potential omitted variable bias. Since the effect of any omitted variables that are correlated with the treatment variable ($DPEP_d$) will be included in the estimated coefficient (β), this might lead to underestimating (overestimating) the size of the impact depending on the direction of the bias. For example, if it were the case that children who belong to relatively wealthier households are part of the treatment group, then this would lead to a positive bias on the estimates of β.

Typically, an RD framework can be interpreted as an Instrumental Variable (IV) setup, where the allocation rule on either side of the cutoff provides an IV. The programme assignment rule used by DPEP (that provide treatment to districts depending on whether their female literacy levels were more or less than the national average) allows me to construct an IV to estimate the causal intergenerational impacts of the DPEP programme. In this setup, whether a district's female literacy level (in 1991) was above or below this cutoff (39.2 percent) can be viewed as the instrument that shapes whether or not it receives the DPEP programme (the $DPEP_d$ variable in (1)). I create a dummy variable ($BelowAvg_d$) that takes a value of one if the district to which the individual belongs lies below the literacy cutoff (39.2 percent), and takes a value of zero otherwise.

To be a valid instrument for programme participation, the dummy variable ($BelowAvg_d$) needs to satisfy two conditions. The inclusion restriction requires that the potentially endogenous independent variable of interest ($DPEP_d$) be correlated with the instrument ($BelowAvg_d$). In other words, the instrument should be a strong predictor of programme participation. The second condition is the exclusion restriction under which the instrument ($BelowAvg_d$) should impact the outcome only through the instrumented variable ($DPEP_d$), and not through other variables. The inclusion restriction can be directly tested and I present these results later in the paper.

The exclusion restriction is not directly testable, but I argue that it is likely to have been satisfied in this setting. In my knowledge, there were no other government programme at that time (or since) that were allocated based on the district female literacy rate. Given that there were no discontinuities in the provision of other government schemes before DPEP, it is unlikely that there would be any apriori discontinuities in outcomes around the female literacy cutoff chosen for DPEP. In addition, through some falsification tests I show that there were no discontinuities in pre-determined covariates. Hence this instrument ($BelowAvg_d$) is unlikely to be correlated with any other covariates around this cutoff. I provide a more detailed discussion in the results section.

The first and second stage equations of this kind of a Two Stage Least Squares (2SLS) approach can be written as:

$$\text{First Stage}: \quad DPEP_d = \alpha BelowAvg_d + \gamma X_{idt} + \nu_{idt} \tag{2}$$

$$\text{Second Stage}: \quad Y_{idt} = \beta_1 DPEP_d + \delta X_{idt} + \epsilon_{idt} \tag{3}$$

This would be considered a global approach to estimating the impact of the programme as it uses all of the data at hand. In such a setup, regular least squares assumptions are imposed and it is common to add higher order polynomials of the running variable (district female literacy rate in this case) as additional covariates in the estimating regression. This method of estimation of the causal effect has appeal because it allows researchers to apply parametric assumptions which allow standard least squares estimation and inference methods to be applicable. But, this approach works best when there is minimal possibility of misspecification bias. Gelman and Imbens [2017] provide a detailed exposition of these methods where they show that such estimation techniques may be suboptimal when inferences are being made about boundary points. Since in an RD setup we compare units on two sides of a cutoff, we can interpret it to be in the realm of boundary inference where the treatment status changes at this boundary point. Gelman and Imbens [2017] remark that in these cases there are three major issues with using polynomials with third or higher degree of the running variable- the point estimates get very noisy, the results could change dramatically with the degree of the polynomial and lastly, and the generated confidence intervals might be inappropriate due to poor coverage. The aforementioned issues imply that using a higher power polynomial may not be suitable when using this approach. Hence, I restrict myself to a quadratic polynomial (power = 2) for specifications that use this global approach to estimation.

More recently, there have been several innovations in the field of =estimation and inference using an RD setup. Most of these tools conduct analysis within a certain bandwidth of the cutoff and hence are classified as local polynomial approaches. I use these innovative and novel data driven techniques which provide more robust inference when analyzing data in an RD setup (Imbens and Kalyanaraman [2012], Calonico et al. [2014], Calonico et al. [2014], Cattaneo and Vazquez-Bare [2016], Calonico et al. [2017]). These methods have several advantages over the parametric higher order polynomial estimation used in many RD applications. I use these local approaches as part of my main empirical specifications and use the aforementioned *global* approach more in the capacity of a robustness check and for the sake of completeness.

Under the newer approaches to estimation in an RD framework, to calculate the impact of the programme, one would have to compare observations that are relatively close on either side of the cutoff. In a regression discontinuity design, individuals on either side of the cutoff are compared under the assumption that they are similar on all characteristics except for a difference in probability of assignment to the treatment group. In some cases, this difference may be *sharp*, where the probability of treatment assignment jumps from zero to one when we move from one side of the cutoff to another. In the case of the DPEP programme, this is not the case.

As with most large scale policies implemented in developing countries, the allocation rule was not followed perfectly, which led to slippages at both sides of the cutoff. Some low literacy districts that were supposed to receive the programme, but did not do so. On the other hand, there were districts with a DFLR higher than the national average that received the programme. This pattern is clearly visible in Figure 13, which shows the distribution of the running variable in the treatment and control districts. Since a district's position relative to the female literacy cutoff does not perfectly predict receipt of the DPEP programme, I use a Fuzzy Regression Discontinuity (FRD) design to estimate the impact of the scheme around the allocation cutoff. This method is not *causal* in the strict sense of the word, but has different interpretations based on the assumptions imposed in estimating the coefficient. The techniques used here calculate a Local Average Treatment Effect (LATE), which is the treatment effect for units in the vicinity of the cutoff. It is important to distinguish this from an Average Treatment Effect (ATE), because this impact may or may not be the same at different parts of the distribution. Hence, there needs to be a certain degree of caution in interpreting these results and in extrapolating them within this context and across different contexts.

In general, in the case of a local estimation of an RD effect a polynomial function is estimated in a neighborhood on both sides of the cutoff. For this purpose a bandwidth (h) is estimated and the neighborhood is defined using this bandwidth : $[\bar{x} - h, \bar{x} + h]$. This choice of bandwidth can be done using different procedures. There are some simple ad-hoc ways of selecting this bandwidth, which have their own sets of advantages and disadvantages. I discuss them later in the text. Here I discuss more objective and data driven methods of selecting these bandwidths. The methods used in this paper to estimate the optimal bandwidths are semi-parametric in nature, which impose minimal structure on the data. These semi-parametric methods mostly trade-off between *efficiency* and *robustness*.[6]

One of the earliest semi-parametric bandwidth selection procedure for RD estimation was introduced by Imbens and Kalyanaraman [2012]. They devised an asymptotically optimal band-

[6] Cattaneo and Vazquez-Bare [2016] explain that "... some methods will be more efficient under the assumptions imposed, but more sensitive to violations of these assumptions, while other methods will be more *robust* to such violations but usually at the cost of some loss of precision.".

width selection procedure under the assumption of squared error loss. This bandwidth depended on the unknown functionals of the data generating process. The algorithm they came up with is completely data-driven and has desirable optimality properties under certain conditions. Under this Mean Squared Error (MSE) optimal bandwidth, the bandwidth can be expressed as follows (Cattaneo and Vazquez-Bare [2016]) :

$$h_{MSE} = C_{MSE}.n^{-1/(2p+3)} \tag{4}$$

where n is the sample size, p is the order of the polynomial chosen for the estimation. The constant C_{mse} depends on a number of factors including the kernel, the polynomial form, the bias and variance of the estimator among others. This constant is unknown and needs to be estimated in order to ascertain the bandwidth (h_{mse}). Imbens and Kalyanaraman [2012] propose a plug in estimator that is based on a reference model to calculate an estimated value of the constant ($\hat{C}_{mse}$). This estimated value is then used to calculate the estimated value of the bandwidth ($\hat{h}_{mse}$). Calonico et al. [2014] improve on this procedure and provide a bandwidth selector that has superior finite sample properties compared to the ones suggested above. Therefore, in addition to being completely data driven and providing a mean squared error optimal bandwidth, this improved bandwidth selection procedure also has desirable small and large sample properties (Cattaneo and Vazquez-Bare [2016]). In the estimations here, I use these improved bandwidth selection procedures in my paper.

One of the downsides of using the bandwidth estimators based on MSE criterion (described above) is that they are invalid for inference procedures. The procedure that they use to balance bias and variance makes inference logically inconsistent (for details refer to Calonico et al. [2014]). This issue in these bandwidth selection procedures implies that the regular confidence interval that they produce are cannot be used for inference. In the limiting case where we assume a zero bias, h_{mse} tends to infinity since C_{mse} is inversely proportional to the bias. This tradeoff implies some corrections need to be made to make inference possible in this framework. Such *robust-bias* corrections have been suggested in Calonico et al. [2016] & Calonico et al. [2017].

When the goal of the estimation is inference, then the MSE optimal estimate and the associated robust bias corrected confidence intervals may not be the optimal ones to employ (Cattaneo and Vazquez-Bare [2016]). It has been shown that the bandwidth that reduces the Coverage Error (CE) of the confidence interval might be more appropriate. It is given by

$$h_{CER} = C_{CER}.n^{-1/(p+3)} \tag{5}$$

where C_{CER} is a constant different from C_{MSE} and needs to be estimated based on the data. The confidence intervals based on this bandwidth (h_{CER}) has been shown to have demonstrably superior properties related to the optimality required for inference procedures. In addition, the bandwidth which minimizes the Coverage Error (CE) is also smaller than the bandwidth which minimizes the Mean Squared Error (MSE). That is, the number of observations used in the estimation using MSE is *larger* than (or equal to) the number of observations used in the estimation using the CE method. Ideally, the CE procedure allows for the estimation of a point estimate and the associated confidence interval, but the point estimate should not be considered. Cattaneo and Vazquez-Bare [2016] suggest that although the CER procedure leads to the estimation of a coefficient, it is not very useful in empirical applications. This is owing to the large degree of variability in these point estimates because of the smaller size of the bandwidth. Even so, I report the point estimates and the associated confidence intervals from these estimations. I mainly focus on the confidence intervals and discuss their relevance in assessing the statistical significance of the variables.

Another approach to bandwidth selection is to use ad-hoc neighborhoods around the cutoff. In empirical applications, the most popular among these is the global approach where the whole data is used in the estimation (described earlier). These generally lead to a larger bandwidth than the data driven processes described above (Cattaneo and Vazquez-Bare [2016]). As is obvious from the way bandwidths are selected in this case, these choices are not data driven. They are subjective and depend on the researcher, thus making them hard to compare. These methods have appeal because they allow researchers to apply parametric assumptions which allow least squares estimation and inference. But, they work best when there is minimal misspecification bias, an issue that has been dealt with in Gelman and Imbens [2017]. This potentially implies that the data driven procedures suggested in the RD literature (Calonico et al. [2014], Imbens and Kalyanaraman [2012], Calonico et al. [2014], Cattaneo and Vazquez-Bare [2016], Calonico et al. [2017]) might be superior to the parametric higher order polynomial estimation done in many RD applications. But, I use the parametric 2SLS regression framework with a quadratic form (of the forcing variable) to show the robustness of my results to such a setup. In addition these ad-hoc neighborhood estimates also serve as a benchmark for the point estimate and helps comparisons with other studies that use this framework.

Apart from the empirical challenges, for the RD design to be valid it is critical that individuals not be able to manipulate their treatment status by systematically positioning themselves on either side of the cutoff. If individuals can choose their own value of the running variable, then they can potentially choose whether or not to be part of the treatment group. This would lead to non-random assignment to treatment, which would complicate the identification of the causal impact of the treatment using an RD approach. Such violations could occur in the DPEP case in two potential ways - if sub-national governments (at the state or district level) were able to choose their treatment status or if individuals were able to affect their treatment status through systematic migration.

It is unlikely that states/districts were able to manipulate their values of the running variable (district female literacy rate) since programme allocation was based on census data, which was collected at an earlier point of time by a central authority in India which is independent of state/district oversight. It is highly likely that the states (or districts) had limited knowledge of the exact decision rule regarding the programme prior to DPEP implementation[7], more so because no other government programmes in the past (or since) appear to have been allocated based on the district-level female literacy rate.

While individuals could potentially have determined their treatment status through systematic migration across districts, I argue that this is unlikely and could not have been large enough to bias the estimates that I identify through my analysis. Firstly, migration across districts in India in the 1990's was fairly low (Topalova [2007]). Secondly, the main reasons for migration in India are marriage and employment. Schooling choice (especially primary or secondary schooling) was not a major reason for migration in India, especially in the rural parts, around the time DPEP was implemented. In terms of migration that is related to seeking enhanced education opportunities, most of it might be expected to be confined to the realms of higher education (high school and beyond). Since the DPEP programme mostly constructed primary, upper primary and secondary

[7] As decisions regarding programme placement were being made by the central government in conjunction with the World Bank and other donors

schools, the case for systematic migration affecting the treatment population seems weak.

5 Data

For this analysis, I use data from multiple sources. I describe each of these below.

Annual Survey of Education Report (ASER): Pratham, an Indian non-profit organization has been conducting annual surveys since 2005 to measure the status of children's schooling and educational outcomes across rural India. These surveys provide cross sectional data on the educational profile of children aged five to 16 spanning the entire country. I use data from the years 2007-2014. For all these years, the format of the test, its administration, the scoring rules and other test-related factors have remained uniform.

Pratham administers tests to children to measure their ability on math, reading (local language) and English. There are four questions of increasing levels of difficulty on each subject. Enumerators record the information on the highest difficulty level question that the students are able to answer correctly. I create a score variable for each child on each test. This variable is measured on a scale from zero to four, depending on the level of question that the child was able to answer. A score of zero implies that the child was not able to answer any question on the test, while a score of four means that the child demonstrated the highest level of proficiency on the test. To illustrate what the scoring captures, on the reading section, a score of zero implies that the child could not read anything; a score of one implies the child could read alphabets; a score of two implies a child could read words; a score of three (four) indicates the child could read a paragraph (story).

The data suggests that a large number of children perform relatively poorly on the tests. The tests are designed to test material covered in grades two or three, but more than 40 percent of the sample reported not being able to read at that level. In fact 15 percent of the sample was not able to read anything at all. Similar patterns are found on the math tests - almost 46 percent of children aged seven to eight years in the main sample could not solve questions that a grade two student should have been able to solve. These statistics imply that a large number of children are lagging behind on the level of proficiency they should have achieved for the grade that they are currently in. Given the large number of students not being able to answer any questions on these tests, I create dummy variables for whether the children are able to answer any questions at all. To go with the measurement of reading/math/English level based on the reported test score, these dummy variables provide a marginal measure of educational achievement.

I also estimate a measure of a child's progression through school by constructing a grade-for-age variable. This variable measures whether a child was held back at school or joined school at a later age than he/she should have. This variable takes a value of one if the child is on track in school and a value of zero otherwise. I also create indicator variables for whether the child has ever enrolled in school and for whether the child had ever dropped out of school. The ASER survey included additional questions in certain years which provide additional insights on the quality of a child's English comprehension. In addition to measuring the ability to read a word or a sentence in English, the children were asked to explain the meaning of the word/sentence. I use this information to create a categorical variable that takes a value of one if the child was able to explain the meaning.

One of the key advantages of using the ASER dataset is it's national coverage. This implies that the sample size of the survey is large and that it provides data on all rural districts in India[8]. In addition, the same test is administered across the country, which facilitates comparison across regions. The ASER data is also unique in that it measures educational achievement at home instead of schools. As a result, the sample includes children who have dropped out of school and children who have never enrolled, in addition to children currently attending school.

District Information System for Education (DISE): This is a government database that provides detailed information on the characteristics of individual schools across India. Data is available on the physical infrastructure and amenities present in each government school, as well as information on teachers, enrolment and other characteristics. This dataset helps track the performance of schools over time and is used by the government to prepare report cards for individual districts and states. Headmasters (principals) of each school are required to fill out a nationally standardized survey instrument every year. This information is subsequently transferred to the cluster officials (at the local administrative level) who verify this information and then pass it further along to the administration at the district level. Here the data is aggregated, digitized and checked before it is made available for public use.

[8] This link provides furhter details on the sampling strategy.

I use the data from this source in several ways. If the school construction programme was effective in creating new schools in treatment districts, there should be a significant increase in the number of schools in these districts. The DISE data, with detailed information on individual schools, allows me to test this hypothesis. One of the innovations of this analysis over others that have examined the DPEP programme in India is the use of the exact timing of programme implementation (obtained from government archives as described below) in each treatment district. Using the DISE data, I verify whether the timing of the programme from the government documents matches the actual construction of schools in the districts. The results from this exercise are discussed later in the text. Finally, I use the DISE dataset to show that at the time the sample children started their schooling there were no significant differences in the quality of schooling environment across districts on either side of the cutoff, which highlights that the impacts on children's learning outcomes is due to the exposure of their mothers to the school construction programme and not because of higher quality of schooling in the districts that had received DPEP.

Archived Government Records: In addition to these survey datasets, I also use archived government records to get detailed information about the timing of implementation of the DPEP programme[9]. I collect data from digital and paper archives to create a comprehensive list of programme districts across various states coupled with information on the exact timing of implementation of the programme across these districts. Although the implementation of DPEP was divided into three phases, there were considerable differences in the timing of programme implementation across districts within the same phase.

Using the archival records, I am able to uncover these differences and use them in my empirical strategy. The timing of the programme across districts is critical to analyzing the impacts of the school construction policy as it is vital to accurately define the beneficiaries of the programme. This will be discussed in greater detail in the empirical strategy section. The process of collating programme implementation data was made more challenging by the fact that in the 1990s there were numerous changes to the administrative boundaries of states and districts in India. In a lot of cases, the documents from the mid to early 1990's made reference to these old political boundaries, which then had to be matched to current geographic boundaries.

[9]Since the DPEP started in the early to mid 1990's, a large amount of the documentation pertaining to the programme was initially not in a digital format. Although, some of the documents have been recently digitized, a large amount of information still exists solely in paper format in governmental archives.

District Level Health Survey (DLHS): This is a household-level survey conducted by the government of India every five years since 1998-1999 to monitor the progress of ongoing health programmes in the country and to furnish statistics on the status of maternal and child health, family planning and other reproductive health indicators. In this analysis, I use data from round three (2007-2008) and round four (2012-2013) of this survey. DLHS-3 contains data from 601 districts across 34 states in India and is representative at the district level. I use these datasets to explore the mechanisms through which the school construction programme could have an intergenerational effect on human capital formation. The variables I construct to study such mediating processes are related to the usage and access of Ante Natal Care (ANC), Post Natal Care (PNC), contraception and other related health care services.

6 Results

As a first step, it is critical to show that there is a significant discontinuity in treatment assignment around the programme cutoff. Figure 1 plots the probability of assignment to treatment against the 1991 district level female literacy rate. The figure clearly illustrates that there is a significant difference in the probability of receiving the programme around the cutoff (39.2 percent). Despite their being a significant difference in probability of treatment assignment, it is possible that because of implementation issues such a difference might not be observed in the actual number of schools constructed as a result of this programme. Therefore it is vital to establish that there is a significant difference in the number of schools constructed in treatment and control districts after the initiation of DPEP.

Table 1 shows the impact of the DPEP programme on the number of schools constructed as estimated using the RD framework. The results indicate that there is a significant positive effect on the number of government schools constructed (nearly 258 schools). There is also a positive and statistically significant higher number of schools constructed (nearly 413 schools) in the typical treatment district as compared to the typical control district. This suggests that while DPEP was only for government schools, it was also accompanied with the construction of a higher number of private schools as well. This provides weak evidence of the crowding-in effects of government expenditure on school construction.

It could be argued that what could shape future outcomes is not the total number of schools but the per capita availability of schools. This might be because of systematic differences in the size of districts across the cutoff due to which an increase in the total number of schools might not lead to an increase in per capita availability of schooling. In the second panel of Table 1, I estimate the impact of DPEP on the per capita availability of schools by using schools per 1000 people as an outcome. Columns 4-6 in Table 1 show that there was a significant increase in the per capita availability of schools in the treatment districts.

As a check of the robustness of these results, I conduct the following falsification exercise - I estimate the impact of the programme on the number of schools that were in these districts before the programme was implemented using a conventional RD setup. Ideally one should not find a statistically significant higher number of schools in the treatment districts because the treatment was allocated to relatively worse off districts. Table 2 shows that the coefficient obtained a result of this exercise is negative and insignificant for the all outcomes that I examine. This shows that the treatment districts were areas where there were potentially fewer number of schools to begin with, but experienced higher school construction during the programme years. Additionally, figures 2-5 show the same results through a graph. Figures 2 & 4 show the discontinuity in number of schools constructed after the programme was implemented.

As discussed earlier, the district level graphs shown in figure 6 provides evidence that the programme was successful in raising the rate of growth of school constructions in the treatment districts. The figure shows that there was a spike in the year-on-year rate of growth of schools in treatment districts after the implementation of the programme. Figure 7 shows that there is no such spike in rate of growth of schools in the control districts around the time of programme implementation. This is similar to the relatively flat trends before the implementation of the programme in the treatment districts. Together with the empirical results in tables 1 & 2, the evidence presented in figures 2 to 7 shows that there were more schools constructed in the programme implementation period in the treatment districts as compared to control districts. Hence, the DPEP programme was successful in creating a discontinuity in the number of schools available around the cutoff for the programme.

Now I discuss the intergenerational impact of the DPEP programme on learning outcomes of children in school more than a decade after it's implementation. As discussed in the empirical strategy section, I use estimators based on two different rules - Mean Squared Error (MSE) minimizing and Coverage Error Rate (CER) minimizing. The CER estimates are preferable as compared to the MSE estimates when the objective is inference - but I present results using both methods. There are other parameters of the model that are to be chosen at the discretion of the researcher. These include the choice of kernel and the order of the polynomial used for estimation. I show the relative robustness of my results to different choices of these parameters.

Recall that the tests are scored in a way that zero indicates failure to provide any correct answers and four indicates complete proficiency. Table 5 presents the results of using CER optimal bandwidth with a quadratic polynomial. The results show that DPDP causes the English test scores of children of beneficiaries to go up by 0.25. The average English score within the sample of children is 1.51, which implies that the average child lies between the level of being able to read an alphabet and word. An impact coefficient of 0.25 implies that DPEP was able to nude the child of a typical program beneficiary towards being able to read a word. Similarly, the coefficient on math is 1.27, which is also significant at the one percent level. The average math ability for the sample in the estimation is 1.68, which means an average child is close to being able to read a two digit number. The programme thus enables children in the next generation to be more than a level higher than the corresponding children in the control group. Similarly, there is a positive effect of the programme on reading ability of the children in the treatment group - the coefficient being 0.48 and it is significant at the one percent level.

In addition to the impact on the intensive margin of test scores, I also estimate the impact of the school construction programme on the extensive margin as well. The DPEP programme has a negative impact on the probability of not answering any questions on the reading (-0.33) and the English (-0.19) tests. Both these effects are significant at the one percent level. The impact on the math test is also in the expected direction but is not statistically significant. Taken together, these results imply that along with having an impact on test scores, the programme also has a significant effect on the probability of answering at least one question on the tests.

Next, I examine the impact of the programme on several other intergenerational educational outcomes –enrolment, school dropout and grade progression. I find that the programme has a positive impact on enrollment (0.12) and a negative impact on child dropout (-0.13). Both these effects are statistically significant. In addition, I calculate a dummy variable ontrack which takes a value of one if the age minus grade of the child is at most six. This sort of measures the consistent progression of the child through school in addition to the timely enrollment in school. I find that the programme has a positive and significant impact (0.04) on grade progression, which implies that the children on mothers impacted by the programme are more likely to be enrolled at the right age and move through grades smoothly.

It is important to consider the bandwidth size used in the estimation of the point estimates and confidence intervals in Table 5. As mentioned earlier, the CER optimal bandwidth provides confidence intervals that are superior for inference. This is in comparison to the MSE optimal estimates, that uses a larger bandwidth than the CER optimal ones. This becomes clearer when comparing the results from Tables 5 and 6, which provide the CER and MES optimal estimates respectively. In each of these tables the estimation pairs the respective optimal bandwidth criterion with the triangular kernel and a polynomial degree two. In general the MSE estimates do not provide confidence intervals that can be directly used for inference and a *robust-bias* correction needs to be made to make inference appropriate. In table 6 I report these robust-bias corrected standard errors. The point estimates (and their significance levels) are similar to the ones obtained in Table 5. The point estimates mostly increase in the MSE optimal case, but they always retain their statistical significance despite the larger bandwidths in Table 6.

As noted earlier, the results in tables 5 and 6 use the local polynomial approach to RD estimation. This approach estimates the causal impact by comparing units *close* to the cutoff within a certain bandwidth classified as estimating the Local Average Treatment Effect (LATE).

6.1 Falsification Checks

In my analysis, I look at the intergenerational impact of the DPEP programme on children who would have potentially started their schooling in the year 2005 or beyond in these treatment districts. However, it is possible that as a result of the DPEP programme, by the year 2005, the programme not only reduced the differences in facilities between the treatment and control districts but actually reversed these differences. If this were the case, then the children of women in the treatment group would have begun their schooling in better schooling districts than their control counterparts. One would then expect to find positive learning impacts among the children of beneficiaries not only because of maternal exposure to the DPEP, but also due to the fact that the children in the treatment group (around the cutoff) are now in better schooling districts. //

If the DPEP programme did have such an effect, the schooling conditions and amenities in treatment districts in 2005 (the earliest school start year among children in the sample) should be better than the schooling conditions in control districts around the cutoff. To verify if this is the case, I search for whether there are any discontinuities in the quality of schooling around the programme allocation cutoff in the year 2005. I use data from both district and school level information to examine this hypothesis. Table 3 presents the RD estimates from using the district level information. The results show that there is no significant difference between the treatment and control districts on a number of different indicators that might signal *quality*.

Panel A in table 3 shows that in 2005 there were no significant differences in the total number of government and private schools at the district level around the cutoff, even when looking at primary schools which were the main focus of DPEP. Panels B and C show that in 2005, there were no treatment-control differences in the availability of different school infrastructural facilities. These are measured in terms of the number of schools in a district with the following characteristics - *pucca* buildings, classrooms, blackboards, girls toilets and drinking water. Panel D presents the results for the number of schools that were provided with grants for school development and enhanced learning, which again fail to show lasting DPEP effects. There are no other differences in the provision of other incentives for learning (uniform and books) across schools around the cutoff. These results imply that by 2005, there seems to be no difference in the quality of schooling infrastructure and other provisions in districts across the programme cutoff.

Now while I examine the differences in schooling conditions across the cutoff at an (aggregated) district level in Table 3, it is possible that these aggregated statistics might be masking critical individual school level differences. To probe whether this is the case, I use school-level data to look at similar differences in *quality* outcomes across schools on either side of the cutoff. Table 4 presents the estimates from this analysis. One way in which enhanced resources may lead to higher levels of learning among children is if they were provided extra tools/amenities that assisted learning in schools. Panel A of Table 4 shows that there are no differences in such schooling level infrastructure provision (book banks and computers).

If DPEP treatment districts had been provided with more funding for supervision and monitoring of schools, there might have been greater accountability leading to enhanced effort from teachers. A similar effect would be noticed if the density of schools increased, causing a higher number of schools to be closer to monitoring centers at the sub-district level (blocks and clusters). Results in Panel B show that neither of these patterns are observable in the data. There was no extra supervision from higher level authorities at the Block Resource Center (BRC) or the Cluster Resource Center (CRC), while the average of distance of schools from these monitoring centers was no different in treatment and control districts.

It is also possible that as a result of the expenditure that the state government had to make for the DPEP, the funds received by the schools through other sources would have been reduced. To check for this, I estimate whether there was a difference in the level of funding received under other programmes (like the total literacy mission) around the cutoff. In Panel C of Table 4, I look at the intensive margin of funding patterns. Here I examine differences at the school level in the level of funding received under different grants as well as the amount of money spent out of the funds that are received. The results show that although funding was greater in the treatment districts, the difference was not statistically significantly different.

Interpreted together, the results in Table 4 show that in 2005 there were no statistically significant differences in school quality across the DPEP programme cutoff. This alleviates concerns about the positive effect on the learning outcomes being driven by superior school quality experienced by children in the treatment group. This further adds credence to the argument that the programme had a positive effect on learning outcomes for children due to their mother's exposure to the DPEP programme, and not due to other factors.

In another validity check, I use the same RD setup as in the main results to see if there are any differences on pre-determined or non-educational outcomes. These are outcome variables that have no relation with the DPEP programme, and hence ideally the programme should have no statistically significant impact on them – child female dummy, mother's age, age of the child, birth year and current year rainfall shock[10]. Instead of choosing the bandwidth level in ad-hoc manner, I estimate the impact on these outcomes using a variety of different bandwidths. The results from this exercise are shown in Figures 19 & 20. I plot the point estimates and their 90 percent

[10]The data on rainfall comes from the University of Delaware dataset on precipitation and air temperature (Matsuura and Willmott [2015])

confidence intervals that are estimated using different bandwidths and kernel functions (Triangle and Epanechnikov). I increase the bandwidth used in steps of 0.05 and show that irrespective of the choice of bandwidth size or the type of kernel function, the point estimate of the impact of the programme on these outcomes always consists of the zero value, showing that DPEP is unlikely to have shaped non-education outcomes.

I also verify whether the setup I use detects any impacts for groups that should have been unexposed to the DPEP programme. Women who were 18 years or older at programme implementation would have missed out on the school expansion under DPEP and so their children would not have benefitted. Figure 16 shows the distribution of the year of birth of the women in this falsification sample. I look to see whether DPEP appears to impact outcomes for this falsification sample, but only look at children born to these women in 2000 or later (Figure 17) to be consistent with the age range of the children in the main estimation sample.

In figure 18, I plot the coefficient and 90 percent confidence interval change for different bandwidths. The blue (bold) line presents the point estimates from RD estimations using different bandwidths while the red (dashed) lines represent the respective confidence intervals. The results indicate that the school construction programme had no statistically significant impact on this outcomes for women who were likely to have left school or aged out of the school going age range by the time DPEP was implemented in their districts. This falsification result is in line with what one would expect for a valid RD setup- the programme did not have an impact on kids of women who were older at the time of programme initiation.

6.2 Robustness Checks

In addition to the verification and falsification checks that I conduct in the previous subsection, I check for the robustness of the results to other modifications to the empirical strategy.

Analyzing an RD setup using a local polynomial approach critically depends on certain parameters - the bandwidth, the power of the chosen polynomial form and the kernel function used to assign weights to the observations around the cutoff. Although I use data driven procedures to select the bandwidth used in the RD estimation, the kernel function and the polynomial are to be inputted by the researcher. Hence, it is possible that the impact estimates are sensitive to the type of kernel function and the order of the polynomial that is chosen. First of all, I check how my results vary when the kernel function used in the estimation is changed. Ideally, the results should not be extremely sensitive to the kernel function used for point estimation. In the results in Tables 5 and 6, I use a triangular kernel. As a robustness check in table 8, I use an epanechnikov kernel instead of a triangular one. While the point estimates of the impact of the programme on different outcomes change marginally with this change in kernel function, the coefficients stay significant and there is no change in the direction of the impact.

As mentioned earlier, the order of the local polynomial function used to estimate the RD impact is also specified by the researcher. Hence in another robustness check of the results, I change the order of the polynomial used in the estimation. Since in table 5 I use a polynomial of order 2, I check the stability of the results with a linear polynomial (Table 7) around the cutoff. The results in the different panels of this table show that the overall pattern of the results remains stable in response to this change in polynomial form. As expected, there is are minor differences in point estimates and the significance levels from the results in tables 5 & 6.

As discussed earlier, the local polynomial approach has certain advantages over the global approach of estimating impacts using a RD setting. But, the global polynomial approach, which imposes a least squares assumption, is popular in empirical settings due to the ease of inference. Therefore, I re-estimate the impact of the programme using the global polynomial approach. As pointed out earlier, Gelman and Imbens [2017] discuss that the use of higher order polynomials (like cubic or quartic functions) of the forcing variable in these estimations might lead to noisy estimates and issues with the confidence intervals. I restrict myself to the use of a quadratic polynomial in these estimations. The results are presented in tables 9, 10 & 11. Since the sample and method used to estimate the coefficients is different, one would expect the point estimates with the global approach to differ from those obtained using the local polynomial approach. But what is critical is that the overall pattern of results does not change – that the programme has a positive impact on children's reading, math and English test scores while reducing the chances of not being able to answer any questions on these tests.

I next alter the definition I use for the sample for the analysis. I check the impact of altering f the different criteria I use to define the study sample. First, I alter the age cutoff that I use to define whether mothers were exposed to the programme or not. In the main analysis I assume that women who were 14 years or below at the time of programme implementation would be part of my treatment group. It is plausible to argue that since the school construction mostly impacted primary school children, the age cutoff to define the treatment group should be lower. I re-estimate the results with a more restricted sample- women who were below 12 years of age at the time of programme implementation. The results from this exercise are presented in tables 14 and 15. The coefficients fall in magnitude, but retain their statistical significance for most outcomes. This implies that the results mostly remain stable for a more conservative definition of the treatment group in terms of age restrictions.

In the main specifications, I assign controls districts within a state the average start year of the treatment districts within that particular state. As an alternative, I assign the start year for the control districts using the average start year of treatment districts nationwide. This alternative definition does not take into consideration the potential inter state differences in timing of treatment - which I leverage in defining the main sample. The results from this change are presented in tables 17 and 18. The coefficient estimates are mostly greater than the ones in tables 9-11. But importantly, the effects are in the expected direction and retain their statistical significance.

In the latter approach, I use the school construction data from DISE to determine the start date of the programme in treatment districts. I assign the year with the highest yearly rate of growth of schools in the post 1993 period as the year when the DPEP programme was implemented. The intuition behind this is that the year when the programme took force in a particular district should ideally be the year when the DISE data reflects the maximum yearly rate of growth in school construction. The control districts are given the average start year of the treatment districts in that particular state. The results from these two different definitions of the treatment groups are presented in tables 20- 22 and 23-25 respectively. The programme continues to have positive impacts on the intensive and extensive margin of test scores, and the effects remain statistically significant.

The results in this section clearly show that the DPEP had a positive impact on the learning outcomes of children in treatment districts whose mothers were exposed to the programme. In addition, I show that the children in treatment districts whose mothers were too old to benefit from this programme experienced no effects of DPEP.The programme did not have an impact on unrelated or pre-determined variables. Importantly, I show that there were no significant differences in school quality across the programme cutoff in treatment and control districts in the year

2005 - the earliest year of starting school amongst children in the sample.

I also show the robustness of the results to alternative methods (2SLS IV) of estimating the Local Average Treatment Effect (LATE). I find that the results are robust when I use alternative specifications and other definitions of the treatment and control group. This increases the strength and credibility of my main results.

7 Mechanisms

[11] There are numerous mechanisms through which a school building programme could have impacted intergenerational human capital formation in a developing country like India. Using multiple household survey datasets, I examine the pathways through which this effect could have been mediated. A majority of the mechanisms examined here are based on the effect that the school construction programme had on the mothers of the children for whom I examine outcomes. These were the women who were in school at the time of DPEP implementation.

I find that the mother's education and literacy does go up as a result of the DPEP programme. I also find that the programme had a significant and positive impact on mother's health as measured by height (in centimeters), BMI and hemoglobin (Sunder [2017]). This effect is similar in nature to the effect found by other studies looking at the impact of women's education on their own health (Grossman [2015], Grépin and Bharadwaj [2015], Agüero and Bharadwaj [2014], Lundborg [2013], Amin et al. [2013], Silles [2009], Currie and Moretti [2003] among others). But it is important to point out an important difference- In this paper I find an impact of the school construction programme on the health of the women and not directly the impact of women's education on their own health.

As another mechanism, I test whether the women impacted by the school construction programme had superior marriage market outcomes. One might expect that if women spent more time in schools, then they would marry later and have children later in life (Field and Ambrus [2008]; Sunder [2016]). I examine whether such effects prevailed among DPEP female beneficiaries and find that women who were of school going age at the time of programme implementation married about half a year later than women in the control group and that their age at first birth

[11] The tables for this section are available from the author on request

was pushed back by around 3-4 months. I also find that the programme had a positive impact on overall contraceptive knowledge and usage. Health knowledge could be an important mechanism affecting the human capital of the children since it helps in reducing unwanted fertility. This is especially important in a high fertility context like rural India (Kugler and Kumar [2017]). I find that women in the treatment districts report higher usage of contraceptives and for condoms in particular. This is similar in pattern to the findings of other studies in different contexts (Johnston et al. [2015], Andalón et al. [2014]).

A child's human capital is significantly impacted by in-utero conditions (Almond and Currie [2011] and Currie and Vogl [2013] provide a good reviews of this literature). Therefore factors having an impact during pregnancy (first 1000 days) could potentially have a profound long term impact on later life outcomes for children. Towards this end, I examine the impact of the programme on a variety of women's ante natal and post natal care outcomes. Potentially, the more care that mothers seek during and after pregnancy, the better the health and well being of the child. I find a positive effect of the DPEP programme on the amount and nature of ante natal care that women use in their last reported pregnancy. Among these women, I find that women who were impacted by the programme were two to four percent more likely to use any Ante Natal Care. I also find that they have higher number of ANC visits and also have their first ANC visit earlier in their pregnancies. Additionally, they are one to six percent more likely to receive IFA tablets and are two to three percent more likely to actually take them.

Results suggest that the women impacted by this programme are more likely to give birth in a hospital or similar institution. In addition, the children of these women are nine percent less likely to be of low birth weight as per WHO norms (less than 2500 grams). Based on these findings it is clear that the women receive better care when they are pregnant, which likely leads to better outcomes at birth. I find that this carries over into behaviors and practices after birth. I find that women in the treatment group of DPEP are more likely to access some sort of PNC (eight to 10 percent). Also, the mother is 8 to 12 percent more likely to get the child tested after birth and does so sooner after birth. Thus women impacted by DPEP more likely to access both ANC & PNC.

8 Conclusion

In this paper, I use the exogenous variation provided by a national level school building programme in India to estimate the intergenerational impact of school construction on learning outcomes. I use the geographical and cohort variation in the implementation of the programme in a Regression Discontinuity framework to find the impact that the programme had on the educational outcomes of the children of the women who were of school going age at the time of programme implementation. I find that the programme had positive effects on enrolment and grade progression for children whose mothers went to school after the programme was initiated. In addition, the programme had a positive impact on child reading, math and English test scores and reduced school dropout.

I find that the results are robust to variations in the empirical strategy. I also rule out the possibility that these results are driven by potentially higher quality of schooling in the treatment districts, in comparison to the control districts. I subsequently explore the potential mechanisms through which the intergenerational impacts of the school building might have been mediated. I find that the programme had significant impacts on maternal education, health and health behavior.

The complex needs of the labor markets of today implies that the workforce needs to equip themselves with a variety of skills over and above basic literacy and arithmetic skills.These skills are lacking in a large section of the population in many developing countries. Hence, it is important that there be a focus on skill building. This paper provides evidence that school construction programmes can lead to improvement in these basic skills at an intergenerational level.

In resource scarce contexts, parents themselves are often uneducated, and so they may not fully understand the benefits that a good education might confer. This may be further compounded by their inability to supplement and complement the education that the children get in schools. School construction programmes relax this constraint as they widely increase the access of school to the population. When those impacted positively by these programme become parents, they then are better able to contribute to the physical and cognitive development of their children, thus expanding the impact of the programme for future generations, as are indicated by the results of this analysis.

References

Anjali Adukia. Sanitation and education. *American Economic Journal: Applied Economics*, 9(2):23–59, 2017.

Jorge M Agüero and Prashant Bharadwaj. Do the more educated know more about health? evidence from schooling and hiv knowledge in zimbabwe. *Economic Development and Cultural Change*, 62(3):489–517, 2014.

Douglas Almond and Janet Currie. Killing me softly: The fetal origins hypothesis. *The Journal of Economic Perspectives*, 25(3):153–172, 2011.

Vikesh Amin, Jere R Behrman, and Tim D Spector. Does more schooling improve health outcomes and health related behaviors? evidence from uk twins. *Economics of education review*, 35:134–148, 2013.

Mabel Andalón, Jenny Williams, and Michael Grossman. Empowering women: The effect of schooling on young women's knowledge and use of contraception. Technical report, National Bureau of Economic Research, 2014.

Mehtabul Azam and Chan Hang Saing. Assessing the impact of district primary education program in india. *Review of Development Economics*, 2016.

Robert J Barro and Jong Wha Lee. A new data set of educational attainment in the world, 1950–2010. *Journal of development economics*, 104:184–198, 2013.

Francis Bloch, Vijayendra Rao, and Sonalde Desai. Wedding celebrations as conspicuous consumption signaling social status in rural india. *Journal of Human Resources*, 39(3):675–695, 2004.

Dana Burde and Leigh L Linden. Bringing education to afghan girls: A randomized controlled trial of village-based schools. *American Economic Journal: Applied Economics*, 5(3):27–40, 2013.

Lizette Burgers. Background and rationale for school sanitation and hygiene education. *New York: UNICEF*, 2000.

Anne Burrows 1 and Sally Johnson. Girls' experiences of menarche and menstruation. *Journal of Reproductive and Infant Psychology*, 23(3):235–249, 2005.

Sebastian Calonico, Matias D Cattaneo, and Rocio Titiunik. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326, 2014.

Sebastian Calonico, Matias D Cattaneo, and Max H Farrell. Coverage error optimal confidence intervals for regression discontinuity designs. *working paper*, 2016.

Sebastian Calonico, Matias D Cattaneo, and Max H Farrell. On the effect of bias estimation on coverage accuracy in nonparametric inference. *Journal of the American Statistical Association*, (just-accepted), 2017.

Matias D Cattaneo and Gonzalo Vazquez-Bare. The choice of neighborhood in regression discontinuity designs. *Observational Studies*, 2:134Ā146, 2016.

Stephanie Riegg Cellini, Fernando Ferreira, and Jesse Rothstein. The value of school facility investments: Evidence from a dynamic regression discontinuity design. *The Quarterly Journal of Economics*, 125(1):215–261, 2010.

Janet Currie and Enrico Moretti. Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics*, 118(4):1495–1532, 2003.

Janet Currie and Tom Vogl. Early-life health and adult circumstance in developing countries. *Annu. Rev. Econ.*, 5(1):1–36, 2013.

Jacobus De Hoop and Furio C Rosati. Does promoting school attendance reduce child labor? evidence from burkina faso's bright project. *Economics of Education Review*, 39:78–96, 2014.

Esther Duflo. The medium run effects of educational expansion: Evidence from a large school construction program in indonesia. *Journal of Development Economics*, 74(1):163–197, 2004.

Alicia Fentiman, Andrew Hall, and Donald Bundy. School enrolment patterns in rural ghana: a comparative study of the impact of location, gender, age and health on children's access to basic schooling. *Comparative education*, 35(3):331–349, 1999.

Erica Field and Attila Ambrus. Early marriage, age of menarche, and female schooling attainment in bangladesh. *Journal of political Economy*, 116(5):881–930, 2008.

Andrew Gelman and Guido Imbens. Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, (just-accepted), 2017.

Peter Glick and David E Sahn. Early academic performance, grade repetition, and school attainment in senegal: A panel data analysis. *The World Bank Economic Review*, 24(1):93–120, 2010.

Karen A Grépin and Prashant Bharadwaj. Maternal education and child mortality in zimbabwe. *Journal of health economics*, 44:97–117, 2015.

Michael Grossman. The relationship between health and schooling: What's new? Technical report, National Bureau of Economic Research, 2015.

Guido Imbens and Karthik Kalyanaraman. Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, 79(3):933–959, 2012.

J Jalan and E Glinskaya. Small bang for big bucks. *An Evaluation of a Primary School Intervention in India. Centre for Training and Research in Public Finance And Policy, Calcutta*, 2013.

C Chrisler Joan and PhD Carolyn B Zittel. Menarche stories: Reminiscences of college students from lithuania, malaysia, sudan, and the united states. *Health Care for Women International*, 19 (4):303–312, 1998.

David W Johnston, Grace Lordan, Michael A Shields, and Agne Suziedelyte. Education and health knowledge: Evidence from uk compulsory schooling reform. *Social Science & Medicine*, 127:92–100, 2015.

Heidi Kaila, David Sahn, and Naveen Sunder. Early life determinants of cognitive ability: A comparative study on madagascar and senegal. *Working Paper*, 2017.

Harounan Kazianga, Dan Levy, Leigh L Linden, et al. The effects of. *American Economic Journal: Applied Economics*, 5(3):41–62, 2013.

Gaurav Khanna. Large-scale education reform in general equilibrium: Regression discontinuity evidence from india. *unpublished paper, University of Michigan*, 2015.

Jackie Kirk and Marni Sommer. Menstruation and body awareness: linking girls' health with girls' education. *Royal Tropical Institute (KIT), Special on Gender and Health*, pages 1–22, 2006.

Adriana D Kugler and Santosh Kumar. Preference for boys, family size, and educational attainment in india. *Demography*, 54(3):835–859, 2017.

HEMANSHU KUMAR and ROHINI SOMANATHAN. Mapping indian districts across census years, 1971-2001. *Economic and Political Weekly*, pages 69–73, 2009.

Petter Lundborg. The health returns to schooling—what can we learn from twins? *Journal of population economics*, 26(2):673–701, 2013.

Kenji Matsuura and Cort Willmott. Terrestrial air temperature and precipitation: 1900-2014 gridded monthly time series, version 4.01. *University of Delaware.*, 2015.

Christopher A Neilson and Seth D Zimmerman. The effect of school construction on test scores, school enrollment, and home prices. *Journal of Public Economics*, 120:18–31, 2014.

Raghaw Sharan Pandey. Going to scale with education reform: India's district primary education programme 1995-1999. Technical report, Education Reform and Management Publication Series, 2000.

Tomi-Ann Roberts, Jamie L Goldenberg, Cathleen Power, and Tom Pyszczynski. "feminine protection": The effects of menstruation on attitudes towards women. *Psychology of Women Quarterly*, 26(2):131–139, 2002.

Pragya Sharma, Chetna Malhotra, DK Taneja, and Renuka Saha. Problems related to menstruation amongst adolescent girls. *Indian journal of pediatrics*, 75(2):125–129, 2008.

Mary A Silles. The causal effect of education on health: Evidence from the united kingdom. *Economics of Education review*, 28(1):122–128, 2009.

Marni Sommer. Where the education system and women's bodies collide: The social and health impact of girls' experiences of menstruation and schooling in tanzania. *Journal of adolescence*, 33 (4):521–529, 2010.

Naveen Sunder. Marriage age, social status and intergenerational effects in uganda. Technical report, Working Paper, 2016.

Naveen Sunder. Schooling & health : Causal estimates from a school construction programme in india. *Working Paper*, 2017.

Petia Topalova. Trade liberalization, poverty and inequality: Evidence from indian districts. In *Globalization and poverty*, pages 291–336. University of Chicago Press, 2007.

TABLES

Table 1: **School Construction (post DPEP)**

	Total number of schools			Schools per 1000 people		
	All Schools	Govt.	Private	All Schools	Govt.	Private
RD estimate	413.784**	258.121**	146.903**	0.311*	0.215*	0.081**
	(173.372)	(125.634)	(61.868)	(0.160)	(0.131)	(0.036)
Observations	495	495	495	488	488	488

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

Note: Note: This table refers to the schools constructed in a district *during* the DPEP programme years. The RD estimation uses the **triangular kernel** with **local polynomial of order 2**. The bandwidth selection process is the **one common MSE-optimal bandwidth selector** for the RD treatment effect. Data Source: DISE 2014.

Table 2: **School Construction Falsification (pre DPEP)**

	Total number of schools			Schools per 1000 people		
	All Schools	Govt.	Private	All Schools	Govt.	Private
RD estimate	-43.726	-104.219	50.776	-0.029	-0.023	0.005
	(343.952)	(301.564)	(83.198)	(0.283)	(0.278)	(0.029)
Observations	495	495	495	488	488	488

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

Note: Note: This table refers to the schools constructed in a district *before* DPEP programme years. The RD estimation uses the **triangular kernel** with **local polynomial of order 2**. The bandwidth selection process is the **one common MSE-optimal bandwidth selector** for the RD treatment effect. Data Source: DISE 2014.

Table 3: **School "*Quality*" in 2005 - District Level Data**

Panel A : Total Number of Schools

	Total-Gov		Total-Pvt		Primary-Gov		Primary-Pvt	
	MSE	CER	MSE	CER	MSE	CER	MSE	CER
Robust	-0.42	-1.25	-0.01	-0.04	-0.53	-1.32	0.00	-0.02
	(0.67)	(4.07)	(0.05)	(0.16)	(0.87)	(4.30)	(0.04)	(0.14)
Observations	479	479	479	479	479	479	479	479

Panel B : Infrastructure

	Pucca		No Building		Classrooms		Blackboard	
	MSE	CER	MSE	CER	MSE	CER	MSE	CER
Robust	0.08	-0.10	1.64	5.45	-1.95	-11.61	-0.61	-1.55
	(0.10)	(0.76)	(1.64)	(13.47)	(2.78)	(52.49)	(1.05)	(5.28)
Observations	479	479	479	479	479	479	479	479

Panel C : Other Infrastructure

	Common Toilet		Girls Toilet		Drinking Water		SCR 60	
	MSE	CER	MSE	CER	MSE	CER	MSE	CER
Robust	-0.27	-0.83	0.01	-0.09	-0.17	-0.88	-0.06	-0.02
	(0.54)	(2.97)	(0.26)	(0.66)	(0.45)	(3.21)	(0.18)	(0.21)
Observations	479	479	479	479	479	479	479	479

Panel D : Grants & Incentives

	TLM Grant		Dev Grant		Uniform		Textbook	
	MSE	CER	MSE	CER	MSE	CER	MSE	CER
Robust	-0.64	-1.26	-0.53	-0.94	-26.18	-76.97	-63.72	-343.09
	(1.19)	(4.56)	(1.23)	(3.38)	(38.51)	(251.27)	(82.72)	(1470.94)
Observations	479	479	479	479	479	479	479	479

Note: All outcomes are in 1000s of schools. Panel A consists of outcomes on total number of schools of different types. Panel B & C consist of number of schools with those particular infrastructural facilities present in the school. Panel D consist of outcomes related to amount of grants and incentives received. *SCR 60* stands for student classroom ratio above 60, *Uniform/Books* refer to incentives received in the form of uniforms and books. The **triangular kernel** with **local polynomial of order 2** is used to construct the point estimates. Two bandwidth selectors are used - **one common MSE-optimal bandwidth selector** & **one common CER optimal bandwidth**. **Robust bias corrected standard errors** are reported and **clustering is at the district level.** Estimates are based on author's calculations using the district level data from DISE (2005).

Table 4: **School "*Quality*" in 2005 - School Level Data**

Panel A : School Infrastructure

	Working Days		Inspection		Book Bank		Num Computers	
	MSE	CER	MSE	CER	MSE	CER	MSE	CER
Robust	-16.23	-13.67	-3.89	-4.55	0.41	0.40	1.22	1.00
	(21.66)	(28.20)	(14.52)	(10.05)	(4.56)	(1.42)	(4.61)	(1.94)
Observations	1017788	1017788	1017894	1017894	999052	999052	1017894	1017894

Panel B : Inspections

	Visits-BRC		Visits-CRC		Distance-BRC		Distance-CRC	
	MSE	CER	MSE	CER	MSE	CER	MSE	CER
Robust	6.15	-7.09	-8.49	-18.58	9.41	-1.09	8.04	4.46
	(12.01)	(21.42)	(27.56)	(45.71)	(79.54)	(22.30)	(36.63)	(11.27)
Observations	1017894	1017894	1017894	1017894	1017894	1017894	1017894	1017894

Panel C : Grants (in 1000's of Rupees)

	Dev Grant-Recd.		Dev Grant-Spent		TLM Grant-Recd.		TLM Grant-Spent	
	MSE	CER	MSE	CER	MSE	CER	MSE	CER
Robust	3.44	14.00	2.32	15.97	1.00	3.44	2.12	2.59
	(9.87)	(45.24)	(7.82)	(48.50)	(2.30)	(5.26)	(2.09)	(3.46)
Observations	1017894	1017894	1017894	1017894	1017894	1017894	1017894	1017894

Note: *BRC* & *CRC* refer to the Block Resource Center and the Cluster Resource Center respectively. *SCR 60* stands for student classroom ratio above 60, *Uniform/Books* refer to schools receiving incentives in the form of uniforms and books. *TLM Grant* refers to grants received under the Total Literacy Mission. The **triangular kernel** with **local polynomial of order 2** is used to construct the point estimates. Two bandwidth selectors are used - **one common MSE-optimal bandwidth selector** & **one common CER optimal bandwidth**. **Robust bias corrected standard errors** are reported and **clustering is at the district level.** Estimates are based on author's calculations using the individual school level data from DISE (2005).

Table 5: **Impact of DPEP on Children's outcomes**- CER Optimal

	Read Nothing	Math Nothing	English Nothing	Child Dropout	Enrolled
RD Estimate	-0.332***	0.048	-0.189***	-0.055***	0.10***
S.D.	0.083	0.076	0.0212	0.011	0.018
Bandwidth	0.12	0.10	0.06	0.098	0.13
Effective Obs.	50690	45204	16863	44138	51154

	English Score	Read Score	Math Score	On Track
RD Estimate	0.247***	0.4753***	1.269***	0.043***
S.D.	0.0286	0.0452	0.384	0.009
Bandwidth	0.05	0.08	0.05	0.03
Effective Obs.	13159	37203	28240	25604

Note: Sample consists of children born in or after the year 2000 to women below the age of 14 years at the time of programme implementation in the treatment districts. The programme start dates for the treatment districts come from the government archives. The starting year of the programme for control districts is the state average starting year of treatment districts within the same state. The **triangular kernel** with **local polynomial of order 2** is used to construct the point estimates. The bandwidth selection process is the **one common CER-optimal bandwidth selector** for the RD treatment effect. The effective number of observations indicates the number of observations that lie within the bandwidths indicated in the table. Note that these are *not* the full sample sizes. **Clustering is at the district level and robust bias corrected standard errors are reported.** Estimates are based on author's calculations using the ASER data (2007-2014).

Table 6: **Impact of DPEP on Children's outcomes** - MSE Optimal

	Read Nothing	Math Nothing	English Nothing	Child Dropout	Enrolled
RD Estimate	-0.231***	-0.14***	-0.20***	-0.127***	0.09***
S.D.	0.087	0.028	0.0184	0.0211	0.022
Bandwidth	0.283	0.23	0.12	0.22	0.27
Effective Obs.	106012	96374	27448	76278	80078

	English Score	Read Score	Math Score	On Track
RD Estimate	0.5213***	0.59***	1.347***	0.253***
S.D.	0.0548	0.1516	0.218	0.019
Bandwidth	0.09	0.18	0.11	0.36
Effective Obs.	21406	73180	46664	102179

Note: Sample consists of children born in or after the year 2000 to women below the age of 14 years at the time of programme implementation in the treatment districts. The programme start dates for the treatment districts come from the government archives. The starting year of the programme for control districts is the state average starting year of treatment districts within the same state. The **triangular kernel** with **local polynomial of order 2** is used to construct the point estimates. The bandwidth selection process is the **one common MSE-optimal bandwidth selector** for the RD treatment effect. The effective number of observations indicates the number of observations that lie within the bandwidths indicated in the table. Note that these are *not* the full sample sizes. **Clustering is at the district level and robust bias corrected standard errors are reported.** Estimates are based on author's calculations using the ASER data (2007-2014).

Table 9: **Full Sample Regressions - Exact Timing**

	Enrolled	Child Dropout	On Track
RD Estimate	0.002	-0.002	0.024***
	(0.002)	(0.001)	(0.004)
Observations	469762	469595	464340

Table 10: **Full Sample Regressions - Exact Timing**

	Reading	Math	English
RD Estimate	0.185***	0.099***	0.179***
	(0.020)	(0.017)	(0.029)
Observations	486433	484647	252031

Table 11: **Full Sample Regressions - Exact Timing**

	Read Nothing	Math Nothing	English Nothing
RD Estimate	-0.048***	-0.049***	-0.043***
	(0.006)	(0.006)	(0.010)
Observations	486433	484647	252031

Note: The sample for tables 9, 10 and 11 consists of children born in or after 2000 to mothers who were below 14 years of age at the time of implementation of the DPEP programme in their district. This programme implementation timing is derived from detailed government archives that describe the exact process of programme implementation. The corresponding population in the control districts is identified on the basis of the average start date in treatment districts within the same state. Other controls in the regression include state fixed effects, birth year rainfall shocks, current year rainfall shocks along with quadratic polynomial of the running variable. Standard errors are robust and clustered at district level.

Table 7: **Robustness Check - Power 1** - MSERD & CERRD

MSE Optimal	English Score	Read Score	Math Score	Enrolled
RD Estimate	0.47***	0.29***	1.47***	0.063***
S.D.	0.052	0.04	0.35	0.013
Bandwidth	0.05	0.06	0.09	0.06
Effective Obs.	14642	32451	37669	30274
CER Optimal	English Score	Read Score	Math Score	Enrolled
RD Estimate	0.375***	0.47***	1.23***	0.09***
S.D.	0.043	0.048	0.329	0.025
Bandwidth	0.03	0.03	0.237	0.03
Effective Obs.	7140	19780	25603	14756

MSE Optimal	Read Nothing	Math Nothing	English Nothing	Child Dropout
RD Estimate	-0.634***	-0.125	-0.12***	-0.104***
S.D.	0.099	0.086	0.0159	0.017
Bandwidth	0.05	0.03	0.14	0.08
Effective Obs.	28316	37669	29673	32471
CER Optimal	Read Nothing	Math Nothing	English Nothing	Child Dropout
RD Estimate	-0.802***	0.04	-0.2443***	-0.061***
S.D.	0.134	0.088	0.024	0.016
Bandwidth	0.03	0.06	0.007	0.04
Effective Obs.	13624	32370	17851	23486

Note: Sample consists of children born in or after the year 2000 to women below the age of 14 years at the time of programme implementation in the treatment districts. The programme start dates for the treatment districts come from the government archives. The starting year of the programme for control districts is the state average starting year of treatment districts within the same state. The **triangular kernel** with **local polynomial of order 1** is used to construct the point estimates. The bandwidth selection processes used are the **one common CER-optimal bandwidth selector** and the **one common MSE-optimal bandwidth selector** for the RD treatment effect. The effective number of observations indicates the number of observations that lie within the bandwidths indicated in the table. Note that these are *not* the full sample sizes. **Clustering is at the district level and robust bias corrected standard errors are reported.** Estimates are based on author′s calculations using the ASER data (2007-2014).

Table 8: **Robustness Check - Epanechnikov Kernel** - MSERD & CERRD

MSE Optimal	English Score	Read Score	Math Score	Enrolled
RD Estimate	0.33***	0.94***	1.38***	0.07***
S.D.	0.038	0.206	0.173	0.008
Bandwidth	0.097	0.12	0.12	0.42
Effective Obs.	26981	56560	56381	139999
CER Optimal	English Score	Read Score	Math Score	Enrolled
RD Estimate	0.19***	1.19***	0.99***	0.04***
S.D.	0.024	0.23	0.32	0.01
Bandwidth	0.05	0.05	0.06	0.05
Effective Obs.	13791	25660	32370	26469

MSE Optimal	Read Nothing	Math Nothing	English Nothing	Child Dropout
RD Estimate	-0.533***	-0.06***	-0.30***	-0.02
S.D.	0.17	0.0232	0.0987	0.03
Bandwidth	0.2	0.24	0.15	0.67
Effective Obs.	48021	93344	48021	231815
CER Optimal	Read Nothing	Math Nothing	English Nothing	Child Dropout
RD Estimate	-0.26***	-0.09	-0.12***	-0.047***
S.D.	0.04	0.09	0.014	0.01
Bandwidth	0.06	0.09	0.07	0.05
Effective Obs.	32451	45395	17851	24025

Note: Sample consists of children born in or after the year 2000 to women below the age of 14 years at the time of programme implementation in the treatment districts. The programme start dates for the treatment districts come from the government archives. The starting year of the programme for control districts is the state average starting year of treatment districts within the same state. The **epanechnikov kernel** with **local polynomial of order 1** is used to construct the point estimates. The bandwidth selection processes used are the **one common CER-optimal bandwidth selector** and the **one common MSE-optimal bandwidth selector** for the RD treatment effect. The effective number of observations indicates the number of observations that lie within the bandwidths indicated in the table. Note that these are *not* the full sample sizes. **Clustering is at the district level and robust bias corrected standard errors are reported.** Estimates are based on author′s calculations using the ASER data (2007-2014).

Table 12: **Gender Differences**

	Reading Score		Math Score		English Score	
	Girl	Boy	Girl	Boy	Girl	Boy
RD Estimate	0.199***	0.175***	0.112***	0.086***	0.225***	0.133***
	(0.029)	(0.027)	(0.025)	(0.024)	(0.042)	(0.040)
Observations	230101	253913	229266	252977	119616	130680

Table 13: **Gender Differences**

	Read Nothing		Math Nothing		English Nothing	
	Girl	Boy	Girl	Boy	Girl	Boy
RD Estimate	-0.049***	-0.048***	-0.047***	-0.051***	-0.057***	-0.030**
	(0.009)	(0.009)	(0.009)	(0.008)	(0.015)	(0.014)
Observations	230101	253913	229266	252977	119616	130680

Note: The female and male coefficients come from identical equations run on the female and male samples separately. The sample for tables 12 and 13 consists of children born in or after 2000 to mothers who were below 14 years of age at the time of implementation of the DPEP programme in their district. This programme implementation timing is derived from detailed government archives that describe the exact process of programme implementation. The corresponding population in the control districts is identified on the basis of the statewise average start date in the treatment districts. Other controls in the regression include state fixed effects, birth year rainfall shocks, current year rainfall shocks along with quadratic polynomial of the running variable. Standard errors are robust and clustered at district level.

Table 14: **Table Description-Robust 12 years**

	Reading	Math	English
RD Estimate	0.132***	0.164***	0.137***
	(0.027)	(0.027)	(0.038)
Observations	309775	310074	164351

Table 15: **Table Description-Robust 12 years**

	Read Nothing	Math Nothing	English Nothing
RD Estimate	-0.037***	-0.039***	-0.025*
	(0.009)	(0.008)	(0.014)
Observations	309775	308648	164351

Table 16: **Table Description-Robust 12 years**

	Child Dropout	Enrolled	On Track
RD Estimate	0.001	0.004*	0.019***
	(0.002)	(0.002)	(0.005)
Observations	292273	293935	290635

Note: The sample for tables 14, 15 & 16 consists of children born in or after 2000 to mothers who were below 12 years of age at the time of implementation of the DPEP programme in their district. This programme implementation in treatment districts timing is derived from detailed government archives that describe the exact process of programme implementation. The corresponding population in the control districts is identified on the basis of the average start date of treatment districts within the same state. Other controls in the regression include state fixed effects, birth year rainfall shocks, current year rainfall shocks along with quadratic polynomial of the running variable. Standard errors are robust and clustered at district level.

Table 17: **Table Description - Robustness National Average**

	Reading	Math	English
RD Estimate	0.263***	0.204***	0.280***
	(0.023)	(0.020)	(0.034)
Observations	486264	484428	247489

Table 18: **Table Description - Robustness National Average**

	Read Nothing	Math Nothing	English Nothing
RD Estimate	-0.054***	-0.054***	-0.047***
	(0.007)	(0.007)	(0.011)
Observations	564336	562277	288620

Table 19: **Table Description - Robustness National Average**

	Child Dropout	Enrolled	On Track
RD Estimate	0.004*	0.002	0.028***
	(0.002)	(0.002)	(0.005)
Observations	552653	551834	544759

Note: The sample for tables 17, 18 & 19 consists of children born in or after 2000 to mothers who were below 14 years of age at the time of implementation of the DPEP programme in their district. This programme implementation timing is derived from detailed government archives that describe the exact process of programme implementation. The difference from main results comes in the definition of the control group. The population in the control districts included in the sample are identified on the basis of the *national* average start date in the treatment districts, rather than the *statewise* average in tables 9-11. Other controls in the regression include state fixed effects, birth year rainfall shocks, current year rainfall shocks along with quadratic polynomial of the running variable. Standard errors are robust and clustered at district level.

Table 20: **Table Description - DPEP start in 1993**

	Reading	Math	English
RD Estimate	0.193***	0.204***	0.319***
	(0.016)	(0.014)	(0.023)
Observations	957351	953291	504861

Table 21: **Table Description - DPEP start in 1993**

	Read Nothing	Math Nothing	English Nothing
RD Estimate	-0.048***	-0.056***	-0.081***
	(0.005)	(0.005)	(0.008)
Observations	957351	953291	504861

Table 22: **Table Description - DPEP start in 1993**

	Child Dropout	Enrolled	On Track
RD Estimate	-0.002	0.002*	0.003
	(0.001)	(0.001)	(0.003)
Observations	941250	946991	934148

Note: The sample for tables 20, 21 & 22 consists of children born in or after 2000 to mothers who were below 14 years of age in 1993. I this case the start of the programme is assumed to be 1993-94, the date when the programme was supposed to have begun. This definition of programme timing is consistent with the definition used by Khanna [2015]. Other controls in the regression include state fixed effects, birth year rainfall shocks, current year rainfall shocks along with quadratic polynomial of the running variable. Standard errors are robust and clustered at district level.

Table 23: **Table Description - DISE start Method**

	Reading	Math	English
RD Estimate	0.139***	0.145***	0.248***
	(0.021)	(0.018)	(0.031)
Observations	450933	449304	231758

Table 24: **Table Description - DISE start Method**

	Read Nothing	Math Nothing	English Nothing
RD Estimate	-0.037***	-0.053***	-0.061***
	(0.007)	(0.007)	(0.011)
Observations	450933	449304	231758

Table 25: **Table Description - DISE start Method**

	Child Dropout	Enrolled	On Track
RD Estimate	-0.001	0.006***	-0.002
	(0.001)	(0.002)	(0.004)
Observations	433434	436041	430838

Note: The sample for tables 23, 24 & 25 consists of children born in or after 2000 to mothers who were below 14 years of age in a particular year. This year is the estimated start date of the programme using the DISE dataset - the year with the maximum year on year rate of growth of schools in a particular district after the implementation of the DPEP programme. Other controls in the regression include state fixed effects, birth year rainfall shocks, current year rainfall shocks along with quadratic polynomial of the running variable. Standard errors are robust and clustered at district level.

APPENDIX

Table A1: **Table Description- Robust 18 years**

	Reading	Math	English	Comprehension	Math Total
RD Estimate	0.180***	0.113***	0.171***	0.304***	0.124***
	(0.013)	(0.012)	(0.019)	(0.072)	(0.013)
Observations	1130018	1125909	610417	94894	1125909

Table A2: **Table Description- Robust 18 years**

	Read Nothing	Math Nothing	English Nothing	Meaning-Word	Meaning-Sentence
RD Estimate	-0.040***	-0.042***	-0.039***	0.068***	0.009
	(0.004)	(0.003)	(0.006)	(0.015)	(0.012)
Observations	1130018	1125909	610417	119726	128525

FIGURES

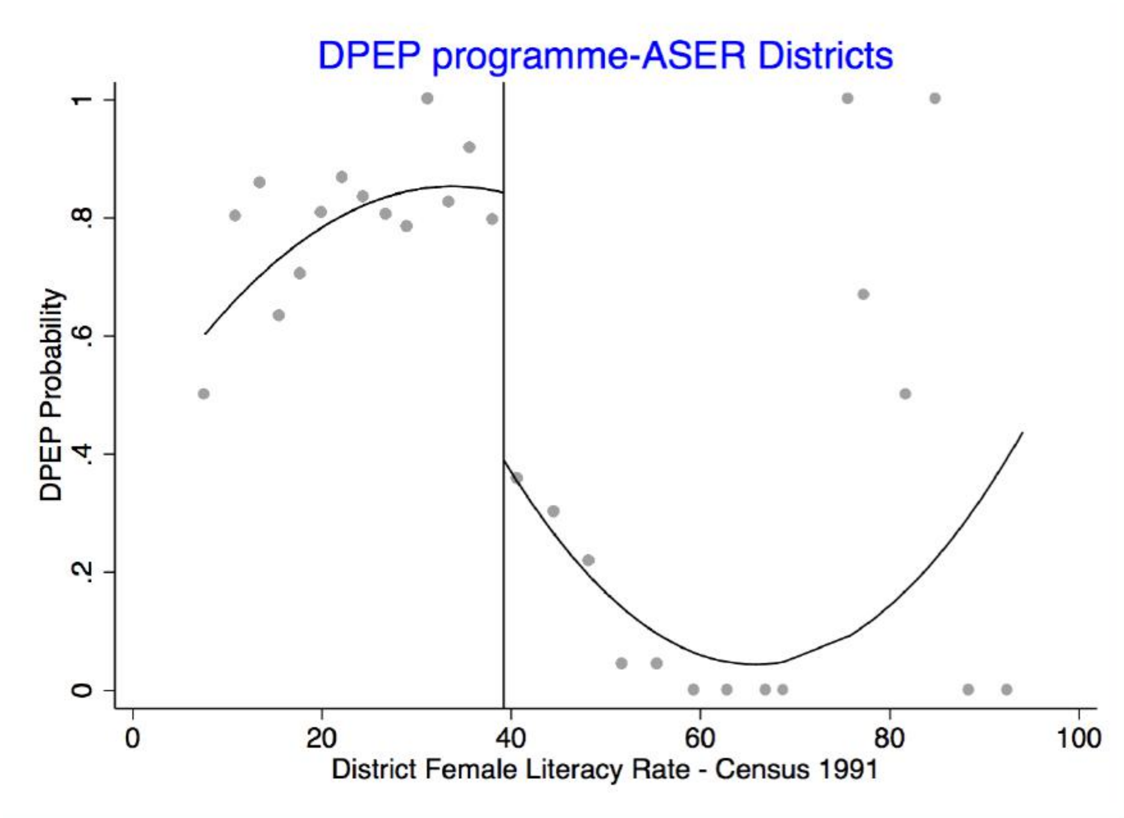

Figure 1: Probability of Receiving DPEP Programme. The graph shows the discontinuity of treatment assignment at the cutoff of 39.2 percent in terms of District Female Literacy Rate. Data Source : ASER data combined with information in government archives

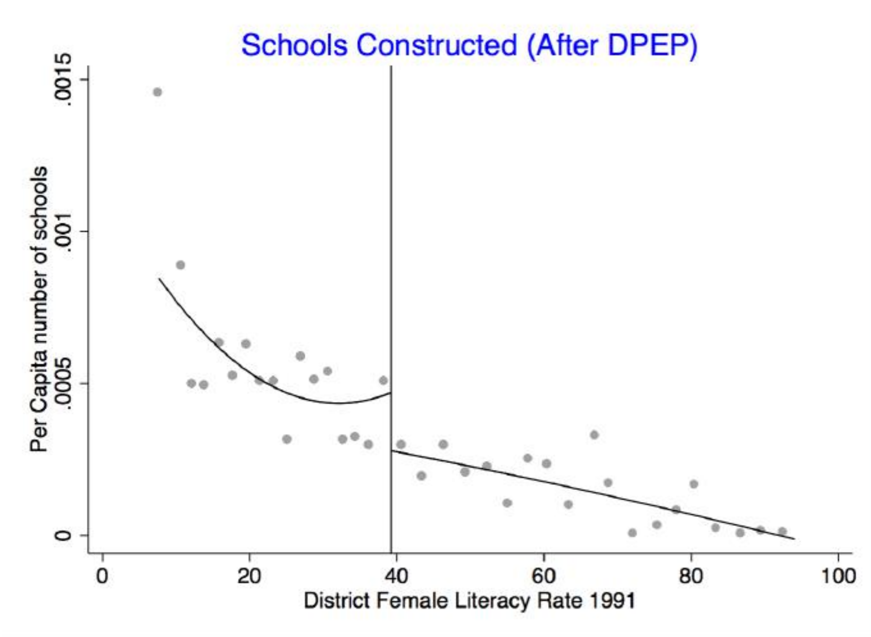

Figure 2: RD Plot : Number of schools constructed (at district level) *after* the implementation of the DPEP programme. Data Source : DISE 2014

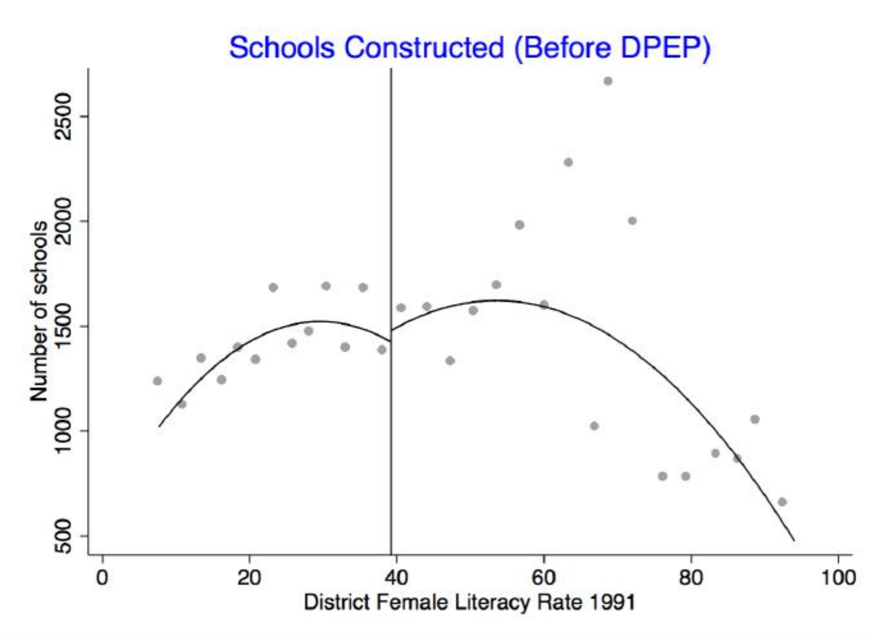

Figure 3: RD Plot : Number of schools constructed (at district level) *before* the implementation of the DPEP programme. Data Source : DISE 2014

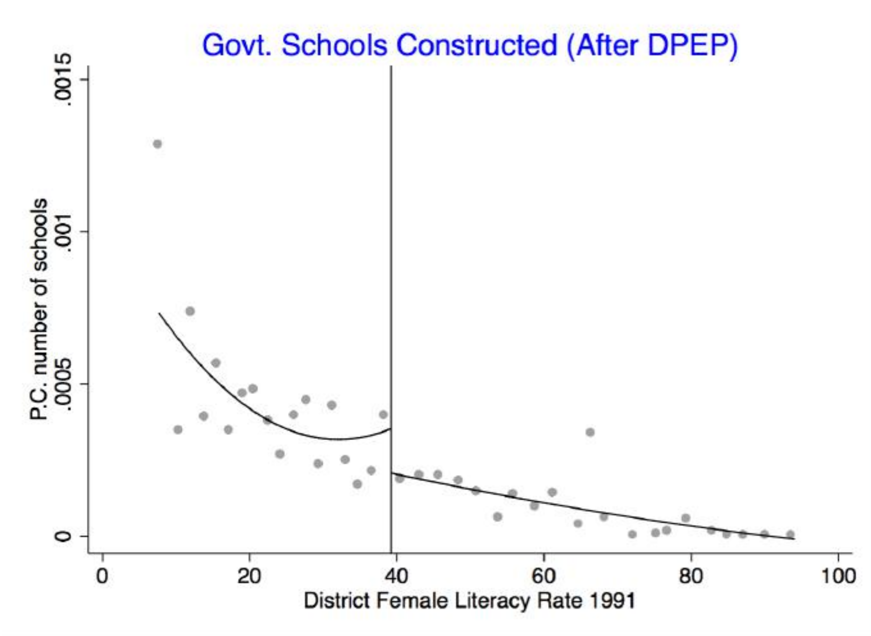

Figure 4: RD Plot : Number of government schools constructed (at district level) *after* the implementation of the DPEP programme. Data Source : DISE 2014

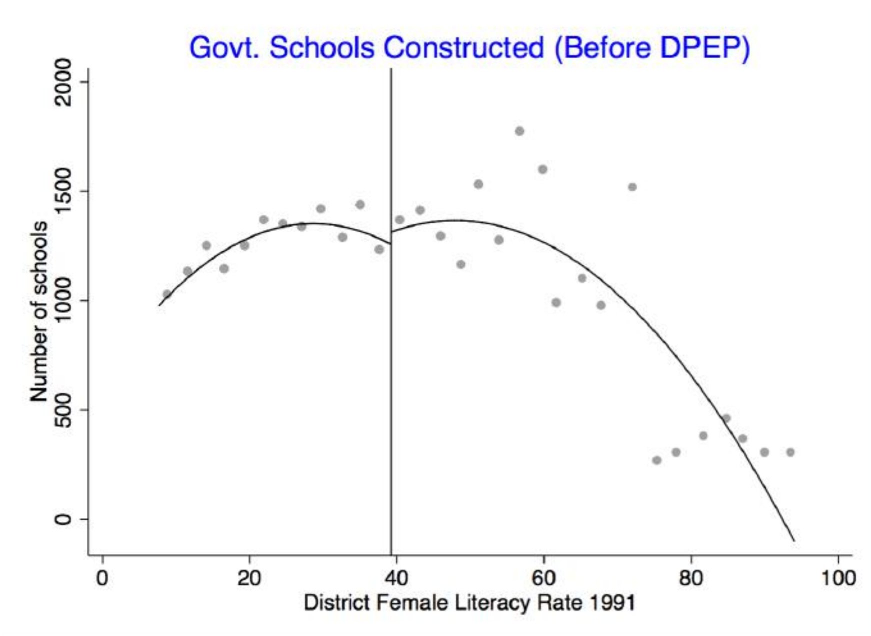

Figure 5: RD Plot : Number of government schools constructed (at district level) *before* the implementation of the DPEP programme. Data Source : DISE 2014

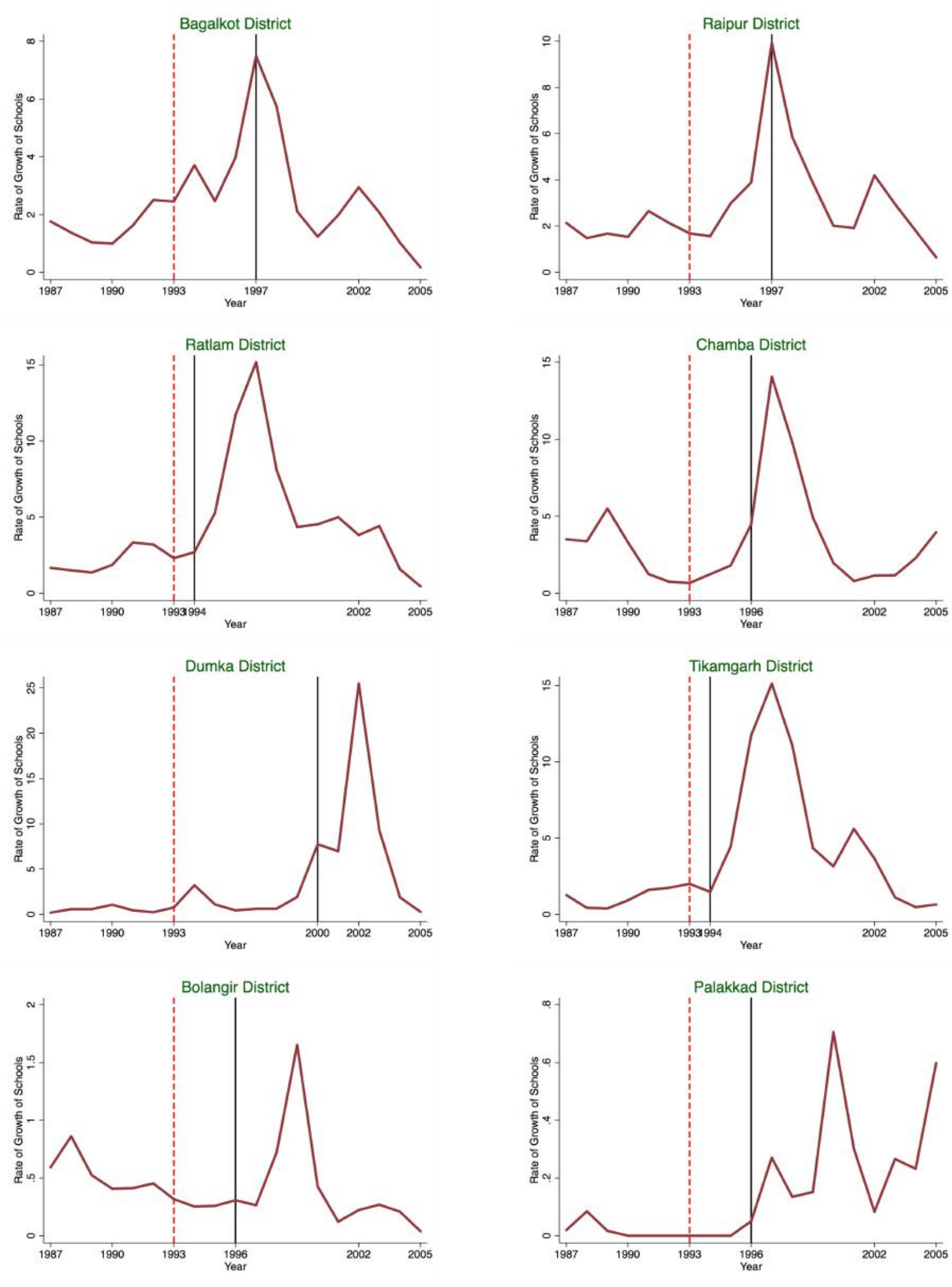

Figure 6: **Yearly rate of growth of school construction plotted against time - Treatment Districts** The graphs in this figure illustrate that the peak in school construction growth in treatment districts is better *predicted* by the year of programme implementation that I infer using the government archival data, rather than the uniform start year of 1993-94. Data : DISE 2005

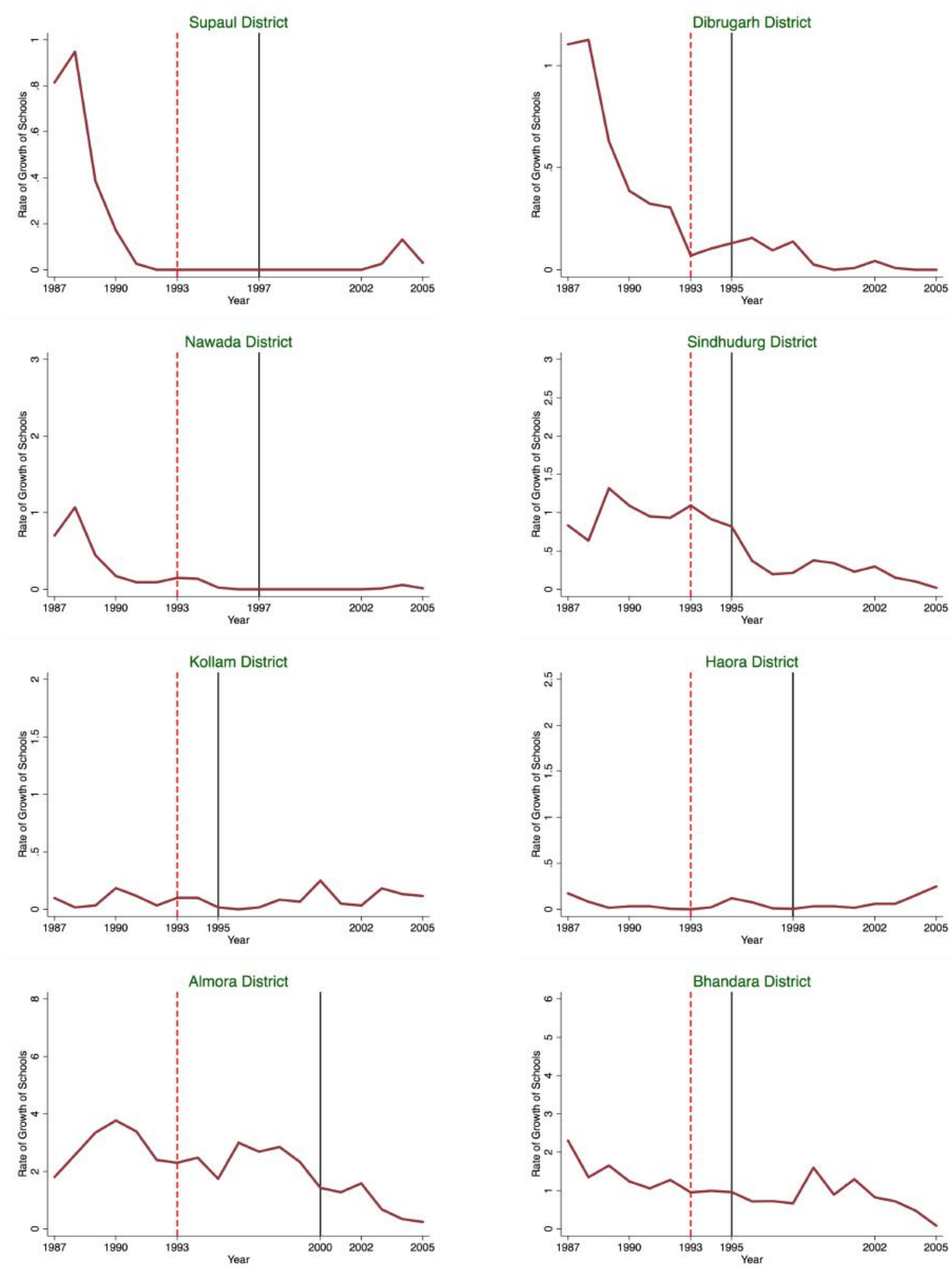

Figure 7: **Yearly rate of growth of school construction plotted against time - Control Districts** The graphs in this figure illustrate that there was no upward trend in school construction in the control districts around the time the DPEP programme was implemented. Data : DISE 2005

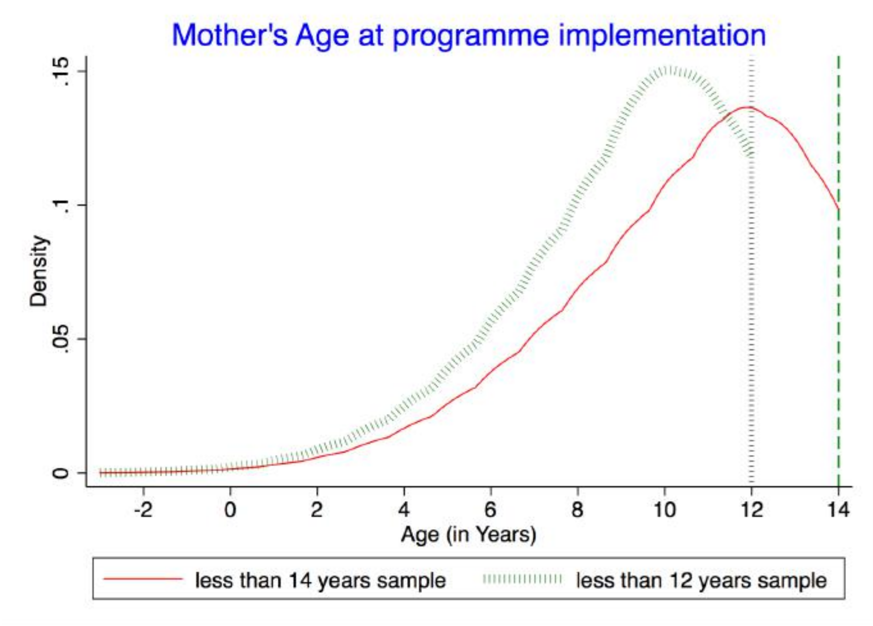

Figure 8: **Age of the mother at the time DPEP implementation in her district.** Here the two curves are based on two different definitions of *at risk of schooling* for women- below 14 years and 12 years respectively at the time of programme implementation. For control districts, the start date assigned is the average start date of the treatment districts within the same state. Based on author's calculations using ASER data from 2007-14.

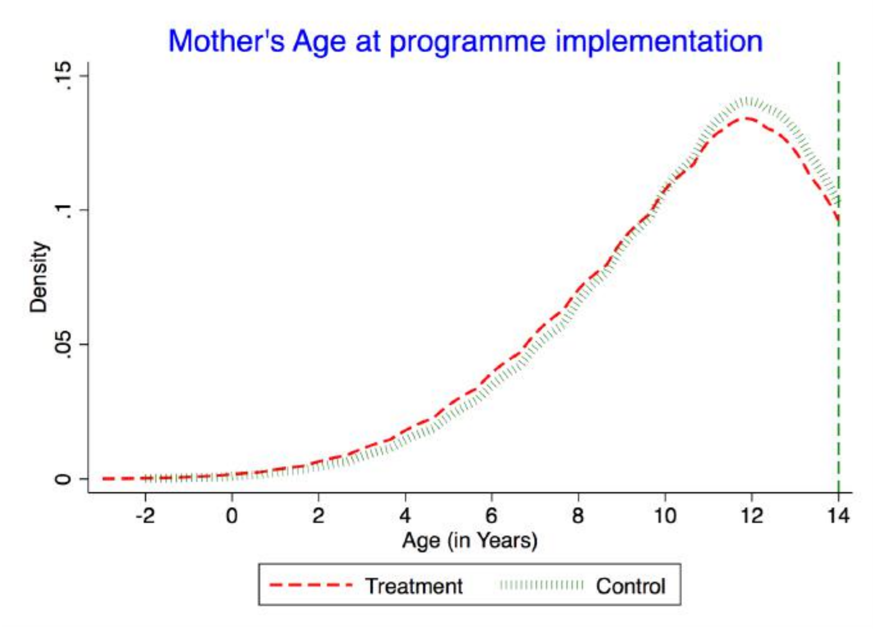

Figure 9: **Age of the mother at the time DPEP implementation across treatment and control districts.** Age of the mother at the time DPEP was implemented in her district. For treatment districts the exact start date is derived from government archival records. For control districts, the start date assigned is the average start date of the treatment districts within the same state. Based on author's calculations using ASER data from 2007-14.

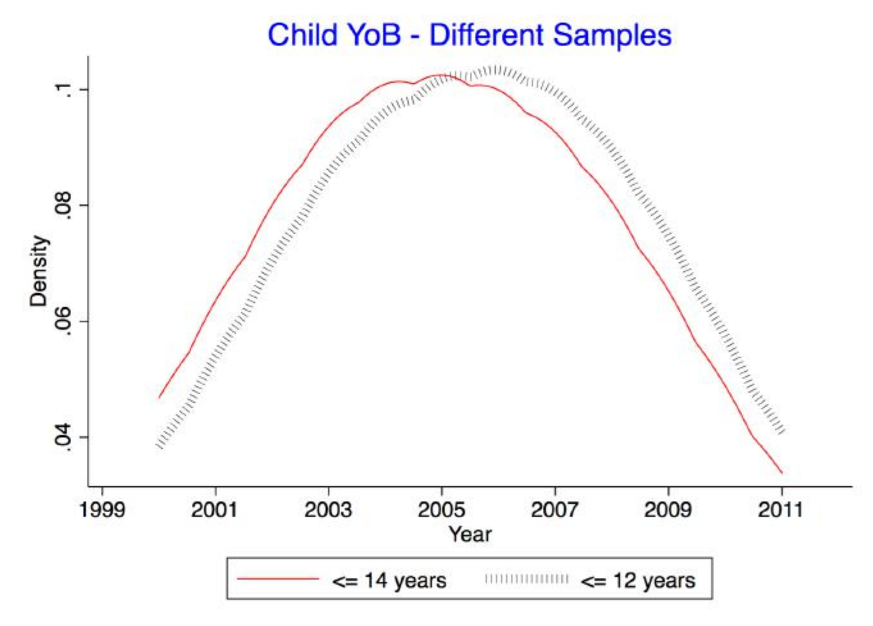

Figure 10: **Year of birth of children in my sample.** Here the two curves are based on two different definitions of *at risk of schooling* for women (mothers)- below 14 years and 12 years respectively at the time of programme implementation. Based on author′s calculations using ASER data from 2007-14.

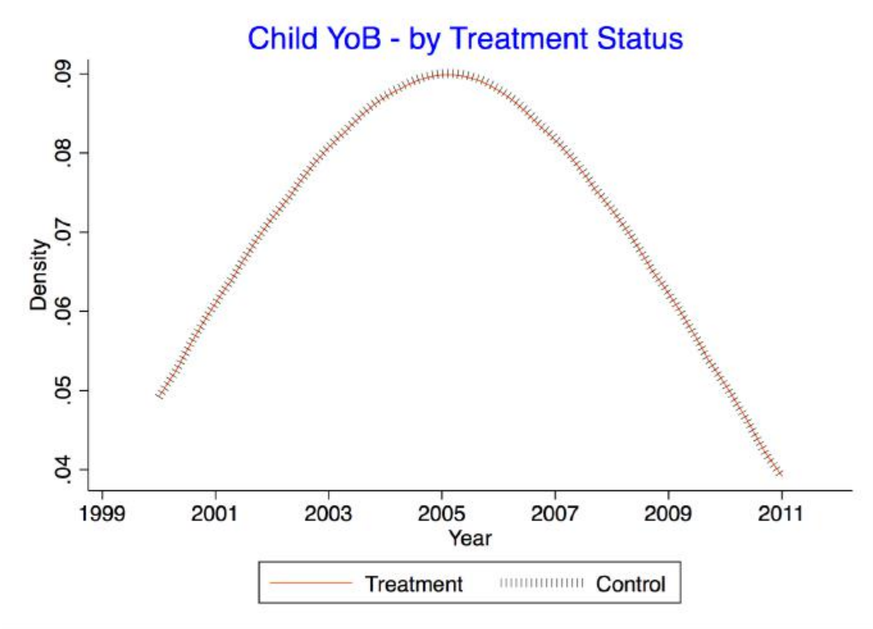

Figure 11: **Year of birth of children in my sample.** The two curves represent the difference in distribution between the treatment and the control groups. Restricted to the sample of children whose mothers were below 14 years of age at the time of programme implementation. Based on author′s calculations using ASER data from 2007-14.

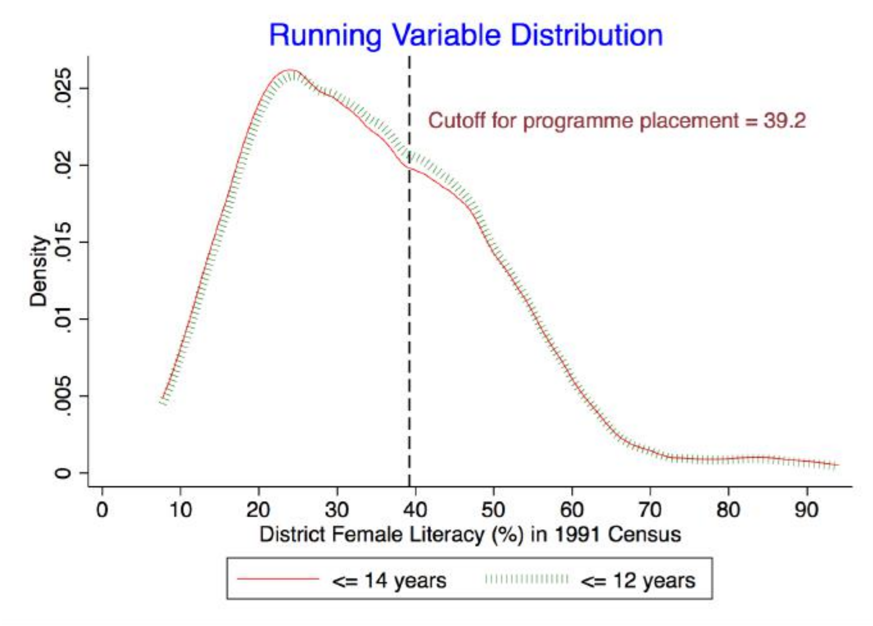

Figure 12: Distribution of the running variable - District Female Literacy Rate (1991 Census). The figure shows that there are no discontinuities in the distribution of the running variable around the programme allocation cutoff. Based on author's calculation using the ASER data from 2007-2014

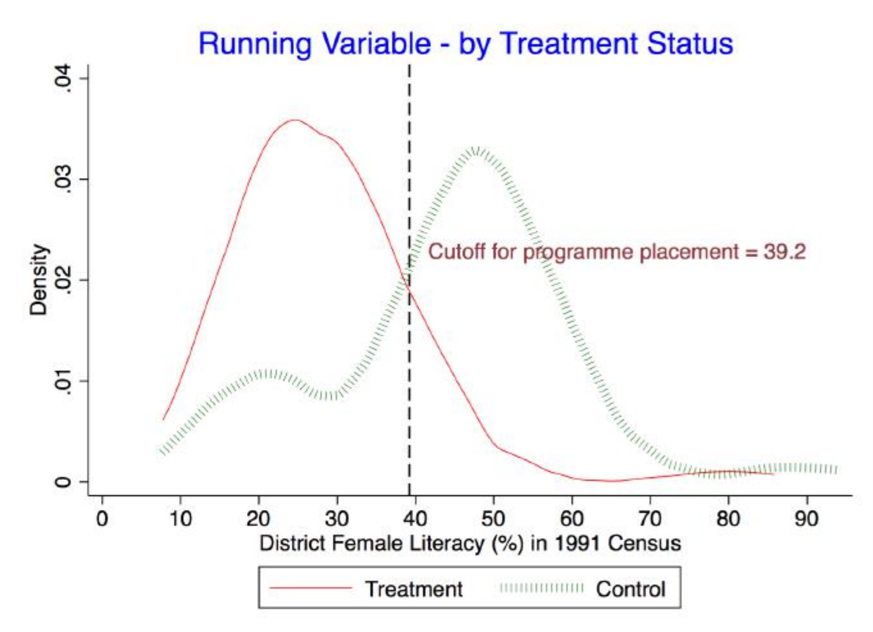

Figure 13: Distribution of the the running variable by treatment status. As expected, the density of treatment districts is higher to the left of the cutoff and the density of the control districts is higher above the cutoff. This is in line with the assignment rule and shows the *fuzzy* nature of the discontinuity around the cutoff.

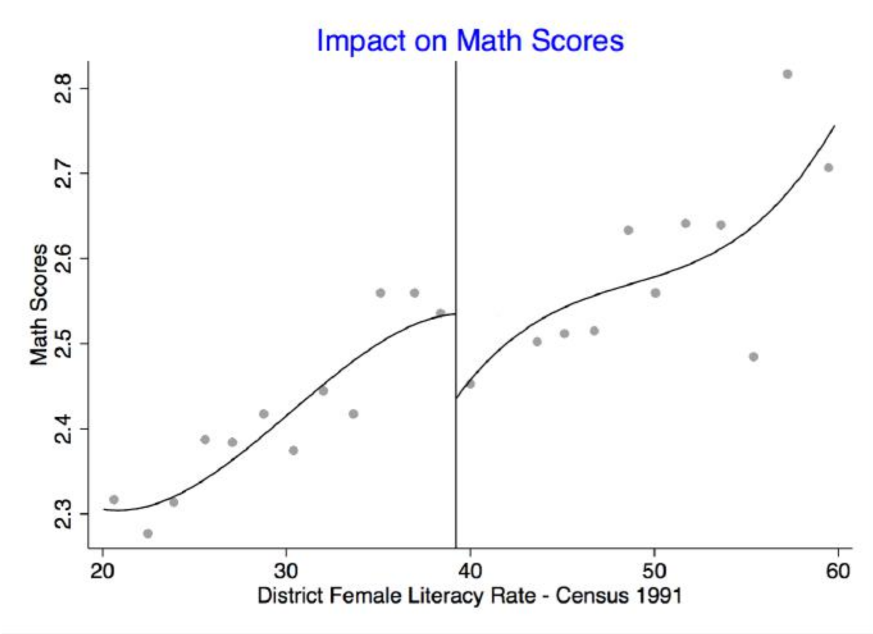

Figure 14: RD Plot illustrating the impact of the DPEP programme on Math Scores. Sample consists of children of women below the age of 14 years at the time of programme implementation in the treatment districts. The programme start dates for the treatment districts come from the government archives. The starting year of the programme for control districts is the state average starting year of treatment districts within the same state. The **triangular kernel** with **local polynomial of order 2** is used to construct the point estimates.

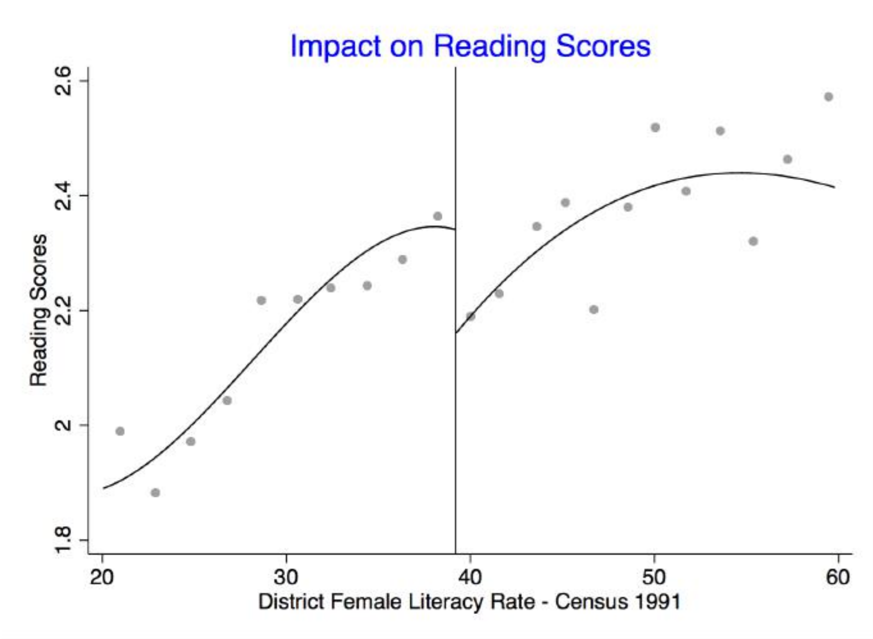

Figure 15: Sample consists of children of women below the age of 14 years at the time of programme implementation in the treatment districts. The programme start years for the treatment districts are inferred from the government archives. The starting year of the programme for control districts is the state average starting year of treatment districts within the same state. The **triangular kernel** with **local polynomial of order 2** is used to construct the point estimates.

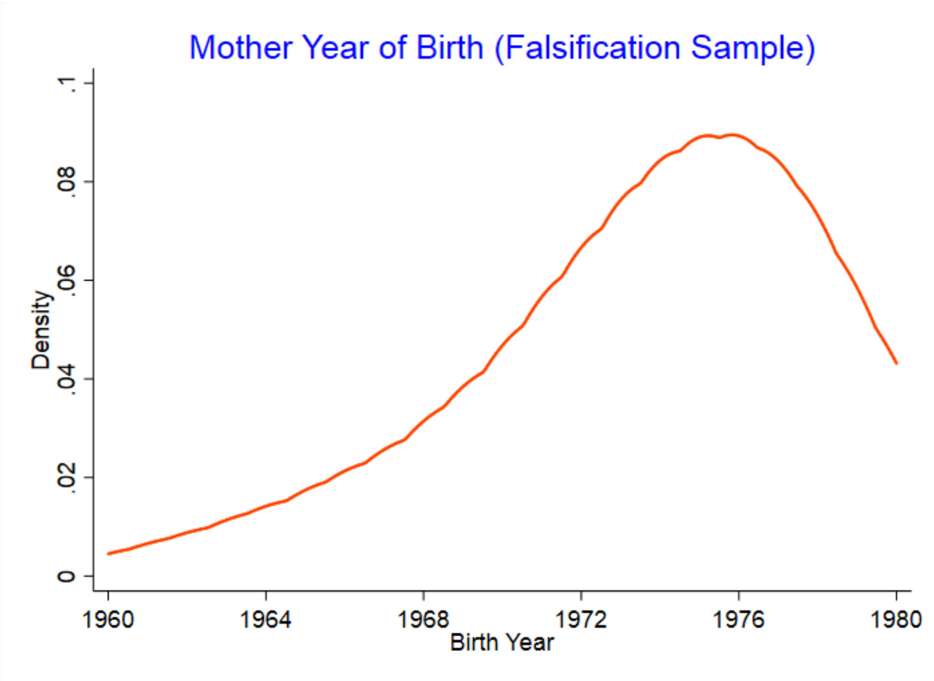

Figure 16: Probability distribution of the Mother's year of birth for the observations in the falsification sample.

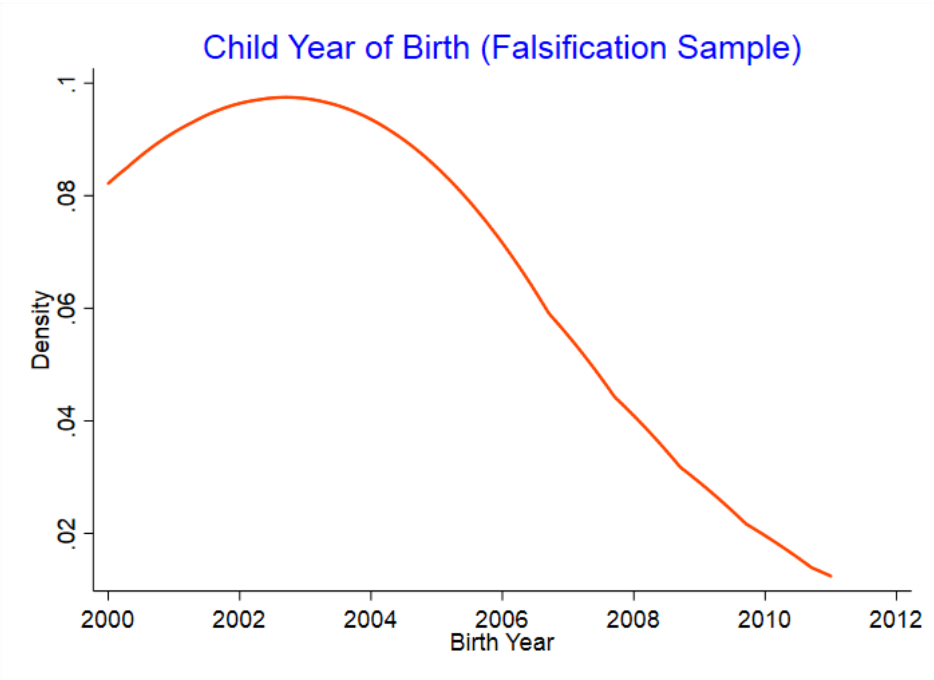

Figure 17: Probability distribution of the Child's year of birth for the observations in the falsification sample.

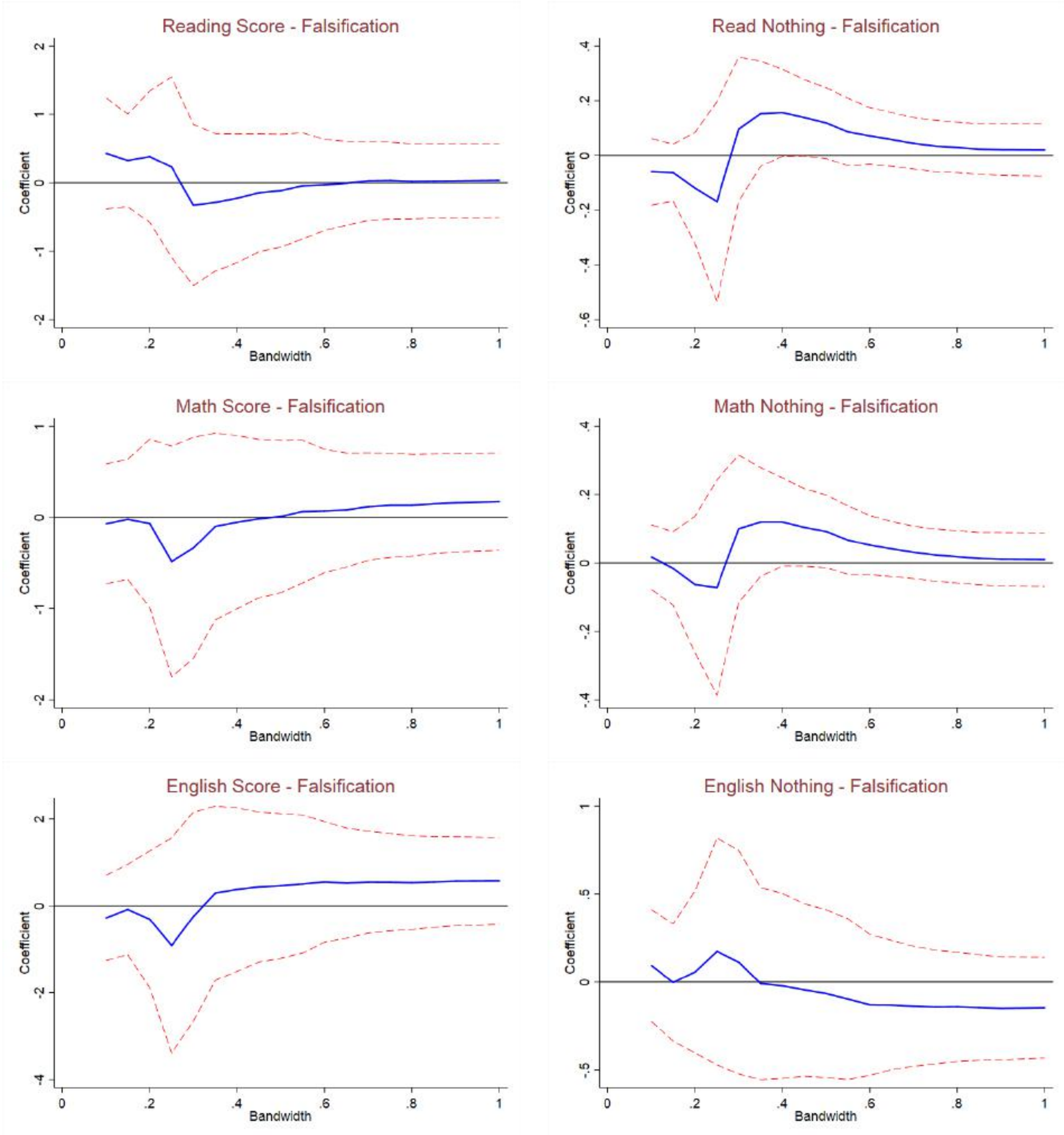

Figure 18: **Falsification Check for sample of Older Women :** These are falsification results for children born in or after the year 2000 to women above the age of 18 years at the time of programme implementation. The staring year for the treatment districts comes from the government archives, while that for the control districts comes from the average starting year of treatment districts within the same state. The point estimates come from RD's that incrementally increase the bandwidth by 0.05 starting from an initial bandwidth of 0.1. Each point estimate comes from a RD setup with a **triangular kernel** combined with a **local polynomial of order 2**. The bandwidth selection process is the **one common MSE-optimal bandwidth selector**. **The robust bias corrected standard errors are used and clustering is at the district level.** Estimates are based on author's calculations using the ASER data (2007-2014).

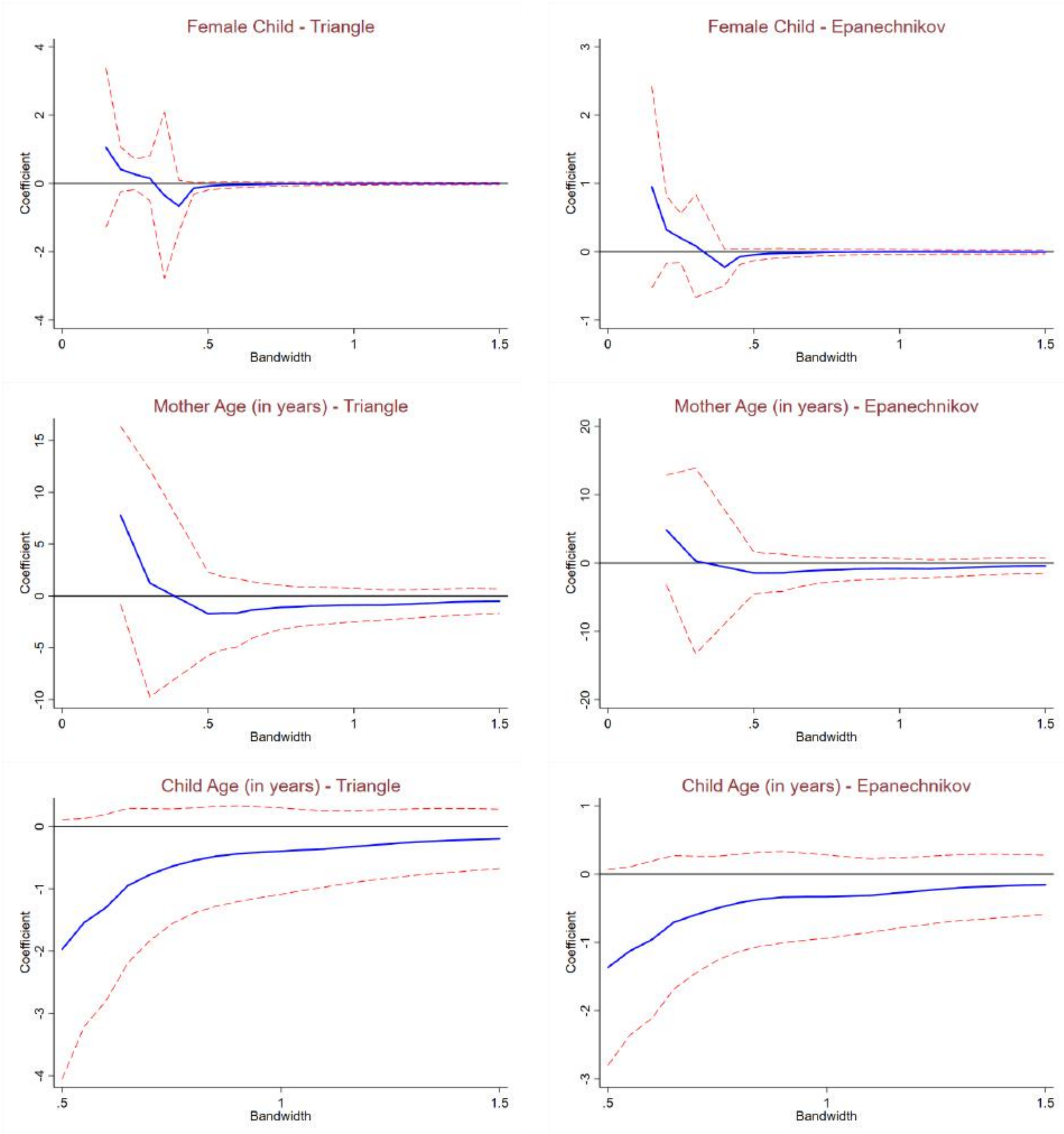

Figure 19: **Discontinuity in Pre-Determined Outcomes** These graphs show that the DPEP programme had a statistically insignificant (i.e. indistinguishable from zero) impact on these outcomes. The staring year for the treatment districts comes from the government archives, while that for the control districts comes from the average starting year of treatment districts within the same state. The left panel shows graphs of the RD impact estimate for an outcome using **triangular** kernel with a **polynomial of degree 2**. The right panel does the same with an epanechnikov kernel. In both graphs the coefficients are estimated at different bandwidths, where bandwidths are increased in steps of 0.05. All estimates use **robust-bias standard errors** with **clustering at the district level**.

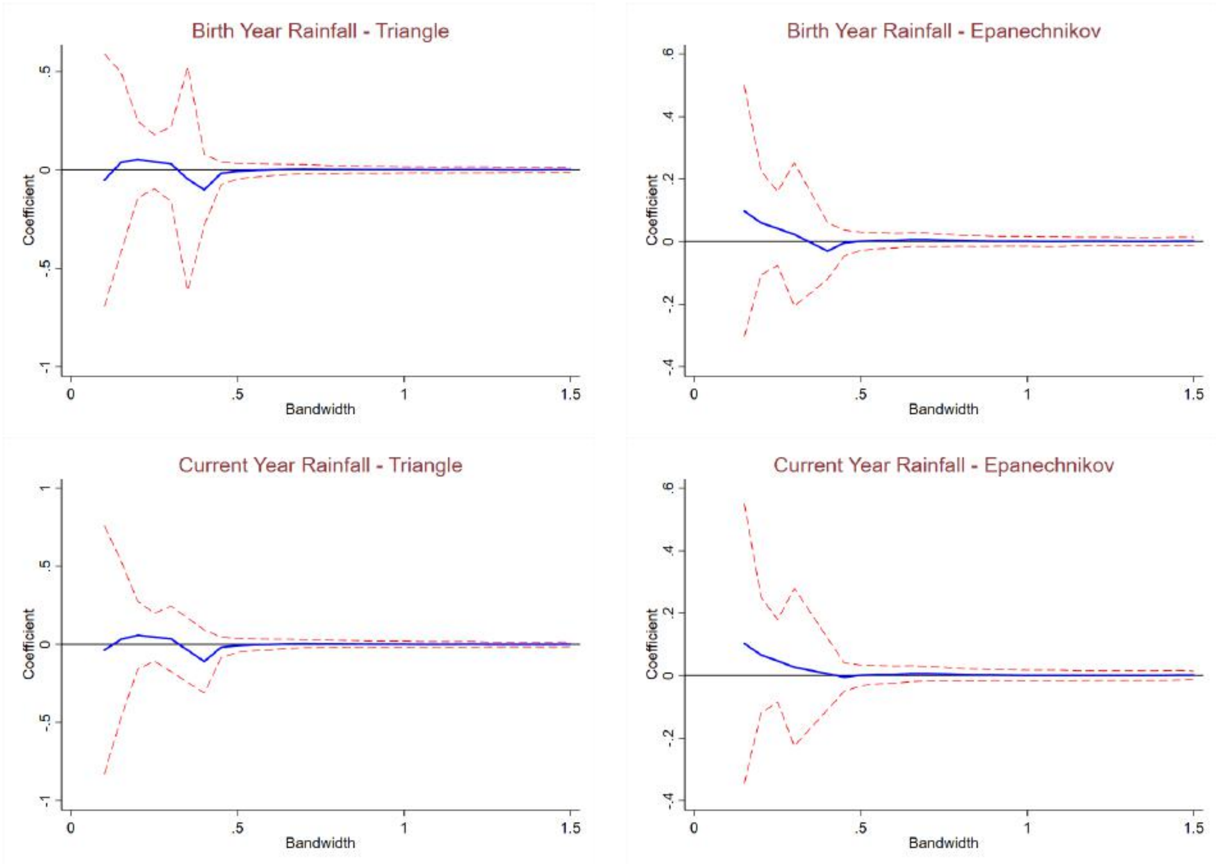

Figure 20: **Discontinuity in Pre-Determined Outcomes** These graphs show that the DPEP programme had a statistically insignificant (i.e. indistinguishable from zero) impact on these outcomes. The staring year for the treatment districts comes from the government archives, while that for the control districts comes from the average starting year of treatment districts within the same state. The left panel shows graphs of the RD impact estimate for an outcome using **triangular** kernel with a **polynomial of degree 2**. The right panel does the same with an epanechnikov kernel. In both graphs the coefficients are estimated at different bandwidths, where bandwidths are increased in steps of 0.05. All estimates use **robust-bias standard errors** with **clustering at the district level**.