School Construction and Intergenerational Human Capital

Naveen Sunder*

Bentley University

April 12, 2025

Abstract

This paper estimates the direct and intergenerational effects of one of the world's largest school expansion policies. Starting in 1993-94, the District Primary Education Programme (DPEP) constructed over 100,000 new schools and served over 50 million children in India. Using a fuzzy regression discontinuity design I show that the policy increased enrollment, literacy and years of education for both male and female direct beneficiaries. Further, I find that children whose mothers were DPEP beneficiaries had higher scores on vernacular (18 percent), math (13 percent), and English (15 percent) tests, while Father's DPEP exposure had no effect on children's learning.

JEL classification: I21, I25, O15, J16, J24

Keywords: Intergenerational Outcomes, Human Capital, India, Regression Discontinuity

*E-mail: *nsunder@bentley.edu*. For their thoughtful discussions and useful comments, I thank Mathew Abraham, Jorge Aguero, Anaka Aiyar, Prateek Bansal, Averi Chakrabarti, Nancy Chau, David Evans, Andrew Goodman-Bacon, George Jakubson, Heidi Kalia, Francesca Marchetta, Prabhu Pingali, David Sahn, Erik Thorbecke, Mallika Thomas and Kira Villa. I also thank seminar participants at Centre for the Study of African Economies (CSAE) 2018, Cornell University, International Association of Applied Econometrics (IAAE) 2018 and the SAU conference 2018. I am grateful to Pratham India for sharing the ASER data with me. The survey datasets used in this article are publicly available – ASER data, District Level Household and Facilities Survey (DLHS), Indian Human Development Survey (IHDS), District Information System for Education. All remaining errors are my own.

Disclosure Statement

This paper has only one author (Naveen Sunder), who has nothing to disclose.

Author Name: Naveen Sunder

Funding Organizations: None.

Support: None.

Position: Assistant Professor, Bentley University

E-mail: nsunder@bentley.edu

Relatives/Partners: Not Applicable.

Review of the Work: Not Applicable.

Institutional Review Board (IRB) Approval: Not Applicable.

Data Availability Statement

Below I list the main survey datasets that are used in this paper - they are all publicly available datasets and can be obtained (or the process is outlined) in the following websites:

- ASER data
- District Level Household and Facilities Survey (DLHS)
- Indian Human Development Survey (IHDS)
- District Information System for Education

The author is willing to assist in procuring these datasets.

1 Introduction

Over the past few decades there has been a sharp rise in the proportion of children enrolled in school, but this has largely been accompanied by mediocre learning levels, especially in developing countries (Barro and Lee, 2013). Learning levels at an early age are critical determinants of a number of long run outcomes. Kindergarten test scores are highly correlated with college attendance and earnings at age 27 (Chetty et al., 2011), while primary and secondary school test scores have an impact on school progression/test performance in high school (Glick and Sahn, 2010), and labor market outcomes (Kaila et al., 2020, Lin et al., 2016). Other evidence highlights the critical role of mathematics scores in shaping a number of labor market outcomes (Glewwe, 1996, Jolliffe, 1998, Moll, 1998, Behrman et al., 2008). These findings are consistent with evidence demonstrating the importance of early childhood investments in determining long term human capital development (Heckman, 2006, 2008)[1].

Although child learning is a function of school, household and individual level factors (Glewwe and Muralidharan, 2016), a bulk of the interventions addressing learning deficits have targeted interventions via schools, and have found encouraging results (see Kremer et al., 2013, Muralidharan, 2013, Muralidharan et al., 2019, Gilligan et al., 2022). In this paper, I focus on the consequences of a household-level factor on child learning outcomes[2] – I study the intergenerational effects of parental schooling on child human capital (Black et al., 2005, Holmlund et al., 2011, Carneiro et al., 2013). Specifically, I evaluate whether parents' exposure to a school construction policy when they were of school-going age impacts their own educational outcomes as well as learning outcomes for their children. The policy I focus on is the District Primary Education Programme (DPEP), which was implemented across 273 districts in 18 Indian states (beginning in

[1] This is also consistent with the macro evidence on the topic: cross-country analyses have also found that cognitive skills are highly associated with economic growth (Hanushek and Woessmann, 2008, 2012).

[2] Earlier studies have examined the effects of different types of household-level interventions on child learning, such as information provision (Jensen, 2010, Loyalka et al., 2013, Wang et al., 2014), conditional cash transfers (Behrman et al., 2009, Baird et al., 2011, Barrera-Osorio et al., 2011), scholarship programmes (Blimpo, 2014, Li et al., 2014) and other in-kind transfers (Oster and Thornton, 2011, Muralidharan and Prakash, 2017).

1993). This programme expanded school access by constructing more than 160,000 primary and upper-primary schools, and remains one of the world's largest primary education programmes in terms of size and funding.

DPEP was targeted towards districts with female literacy below the national average female literacy rate at the time of program initiation, which was 39.2 percent. This assignment rule created a discontinuity in the treatment probability around the threshold of 39.2 percent female literacy – districts just below this cutoff had a much higher probability of receiving DPEP, as compared to districts just above this threshold. This setup enables me to use a fuzzy regression discontinuity (RD) design to estimate the causal impact of this policy on direct beneficiaries (school-going age cohorts during programme years) and their children (intergenerational effects).

I establish that individuals of school-going age (during DPEP years) in targeted districts experienced better educational outcomes than individuals of the same age group in non-DPEP districts. I find that both male and female direct beneficiaries of DPEP had higher enrolment (8-10 percent), literacy (9 percent) and years of education (0.7 - 0.8 years), as compared to non-beneficiaries. Upon exploring the intergenerational effects of the DPEP, I find that children whose mothers were direct beneficiaries of DPEP performed better on vernacular (18 percent), math (13 percent) and English (15 percent) tests, as compared to similar children born to comparable women in non-DPEP districts. Furthermore, the results also suggest that both sons and daughters gain due to maternal DPEP exposure. In contrast to these results, I find small and statistically insignificant effects of father's DPEP exposure on their children (both daughters and sons).

Next, I probe factors that could potentially enable individuals to use their education to impact their children's human capital. I find that women who benefitted directly from DPEP were healthier, had enhanced bargaining power and invested more in their children's education as compared to similar women in non-DPEP districts. Additionally, I conduct a variety of robustness checks to show that the findings are not sensitive to changes in the empirical strategy, variations in the definition of the treatment group and allowances for delays in policy implementation.

This paper contributes to the literature in the several ways. First, I add to the evidence on the critical role played by parents in shaping their children's learning outcomes. Specifically, this study highlights how household inputs could complement school factors and engender better educational outcomes, especially in low/middle income countries. Second, this analysis supplements a growing literature that examines the intergenerational impacts of large-scale schooling expansion policies in developing countries (Akresh et al., 2018 and Mazumder et al., 2021). In doing so, it accounts for the substantial intergenerational gains that could emerge due to these policies, something that had largely been ignored by earlier papers which solely focused on gains realized by direct beneficiaries, thus underestimating the total gains of such programmes. Third, this study is valuable in that it evaluates the learning impacts of an at-scale policy that expanded educational infrastructure. This is uncommon in the literature since papers mostly focus on outcomes such as school enrolment and years of education. Fourth, although this study estimates the local average treatment effect, the findings could be informative for policy formulation in a variety of developing country contexts, since large parts of Central/Western Africa and South Asia have comparable educational indicators to what India had at the time of DPEP implementation. This implies wider policy applicability of the findings of this analysis. India, the context that I study, is a region of the world where such gender-based gaps in years of schooling were particularly pronounced in the 1990s and which, by its sheer scale, accounted for a very large share of women who did not complete primary education at the time. The policy relevance is further enhanced by the fact that girls' education remains a core area for international policy makers.

2 Literature & Background

2.1 Related Literature

There is a large and growing literature on the effects of increasing the supply of schooling infrastructure on educational outcomes in developing countries, the bulk of whcih estimates impacts on direct beneficiaries (Duflo, 2001, 2004, Handa, 2002, Alderman et al., 2003, Burde and Linden, 2013, Kazianga et al., 2013, De Hoop and Rosati, 2014, Giri and Shrestha, 2017). More recently, there have been a couple of papers that examine, like I do in this study, the intergenerational impacts of school construction policies in developing country contexts. Akresh et al. [2018] find that a school construction policy in Indonesia led to better living standards for beneficiaries and had significant intergenerational educational effects for female beneficiaries, but not for male beneficiaries. Another paper examining the same policy finds that the children of female beneficiaries perform significantly better on a primary school national examination (Mazumder et al., 2021). This analysis is closely linked to the latter papers.

This study is also related to the body of literature investigating the general links between parents' education and children's human capital outcomes. A majority of these studies are based on developed countries contexts, but in this paper I add to the growing literature examining these relationships in low-and-middle income countries. This is particularly important as the nature and the extent of these linkages may vary by context. Studies have found that parental education has a direct positive impact on the educational and health outcomes of children (Black et al., 2005, Sacerdote, 2007, Osili and Long, 2008, Chou et al., 2010, Holmlund et al., 2011, Agüero and Ramachandran, 2020)[3]. Children are also likely to benefit indirectly from the improvements that parents experience

[3] This direct link between the educational outcomes of parents and their children has been studied extensively using many different empirical strategies, which include comparing twins and their children (Behrman and Rosenzweig, 2002, Bingley et al., 2005, 2009, Holmlund et al., 2011), natural experiments related to compulsory schooling laws/removal of tuition fees/location (Black et al., 2005, Oreopoulos et al., 2006, Carneiro et al., 2013, Chevalier et al., 2013, Piopiunik, 2014, Agüero and Ramachandran, 2020) and comparing outcomes between biological children and adopted children of the same parents (Sacerdote, 2002, 2007, Silles, 2017). See Björklund and Salvanes [2011] for a review of the associated literature. For a review on the literature related to the impact of parents' education on child health see Grossman [2015].

themselves due to enhanced educational outcomes – improved health (Amin et al., 2013, Agüero and Bharadwaj, 2014, Grossman, 2015, Grépin and Bharadwaj, 2015), higher age at marriage and at first birth (Marchetta and Sahn, 2016, Sunder, 2019), increased contraception usage and better antenatal care practices (Andalón et al., 2014, Johnston et al., 2015, Behrman, 2015), higher bargaining power for women (Lundberg and Pollak, 1993, Duflo, 2012, Samarakoon and Parinduri, 2015) and higher investment in children's education (Andrabi et al., 2012, Yoong, 2012), among other outcomes. In exploring the potential mechanisms that could channel the DPEP's intergenerational effects, I probe several of these factors in the current analysis.

Finally, I add to the fairly scant body of literature that has examined the effects of DPEP till date. Khanna [2021] develops a general equilibrium labor market model, and uses the exogenous variation in education introduced by DPEP to estimate the returns to education, while Azam and Saing [2017] find that DPEP beneficiaries had higher enrollment, total educational attainment and probability of completing primary schooling. In somewhat contrast, Jalan and Glinskaya [2013] find small direct effects on enrolment, concentrated mostly among socially disadvantaged groups. While there are some parallels between the current analysis and these previous DPEP papers, there is one critical difference – the aforementioned papers solely focus on the effects on direct beneficiaries. My results not only confirm that the program had positive impacts on the educational attainment of direct beneficiaries, but also establish that it had significant intergenerational effects, a novel contribution to the literature.

2.2 DPEP Programme

The District Primary Education Programme (DPEP) was the flagship education policy of the Indian government in the 1990s. Implemented in a phased manner across the country, the programme was oriented towards achieving universal education through an increase in schooling infrastructure. Program rules stipulated that DPEP school construction would be targeted towards districts with Female Literacy Rates (the DFLR) below the national average. The programme cutoff was chosen to be the average female literacy in the country at the time, which was 39.2 percent based on the most recently available census data at the time (the 1991 Census). In addition, the central government also specified that DPEP would only be introduced in districts that had successfully implemented the Total Literacy Campaign (TLC), a programme that aimed at improving literacy levels across the country (Rao, 1993). Since the TLC had been implemented successfully across all districts in India by 1994, this criterion turned out to not matter for selection into the DPEP (Jalan and Glinskaya, 2013).

The policy built upon the already existing education infrastructure in the country, and channeled funding and technical assistance to state and district governments for expanding access to primary education. The program was primarily financed by the central government, which bore 85 percent of all costs, for which they received support from donors like Official Development Assistance (ODA), the Royal Government of the Netherlands, the U.K. Department for International Development (DFID) and the World Bank[4]. The remaining 15 percent of the budget was met by state governments. To minimize crowd-out of states' education expenditures, the central government stipulated that states had to continue meeting their existing non-DPEP educational expenditures. Since the DPEP funds were committed over and above the existing education budget, the programme represented a massive surge in education spending across the country. Between 1993 and 2002 public expenditure on elementary education in India rose from 1.7 to 2.1 percent of GDP, accounting for more than 60 percent of the growth in national public expenditure on education for this period (Wu et al., 2005).

The DPEP was initially introduced in 42 districts, and subsequently extended to 273 districts across 18 states. The geographic spread of the policy is illustrated in figure 1[5]. The DPEP programme was phased out starting in 2001-02, at which time a different national program called Sarva Shiksha Abhiyan (SSA) was introduced[6]. This new program

[4]The World Bank alone sanctioned loans worth approximately USD 850 million for this policy alone (Bank, 2007, Kumar et al., 2001).

[5]The DPEP was implemented in phases and hence I aggregate the information related to its implementation from several different sources including government of India reports, state government reports (since education is handled at both federal and state levels in India), non-governmental sources (like World Bank reports – since the World Bank was an implementation partner) and other miscellaneous sources.

[6]Although the DPEP was started to be phased out in 2001-02, there was still non-negative expenditures

was different from DPEP many of its features, but there is one difference that is critical for this analysis – SSA aimed at expanding educational opportunities across the entire country, and was not targeted towards specific districts, like the DPEP.

3 Data

Below, I describe the different data sources that are used in this analysis.

3.1 The Annual Survey of Education Report (ASER)

ASER is an annual national survey conducted to measure schooling and learning outcomes across rural India. These surveys provide repeated cross sectional data on the educational profile of children aged 5 to 16 years. In my analysis, I use these ten rounds of ASER data (years 2007-2014, 2016 and 2018), spanning 21 states. ASER data contain student test scores for math, vernacular and English. A typical ASER test for a subject includes four questions, with each question being more difficult than the previous one. The score assigned to a child indicates the highest level of difficulty that a child was able to solve. For example, on the reading section, a child gets a score of zero if she is unable read anything; one if she is able to recognize alphabets, and two, three or four depending on whether she is at best able to read words, sentences or a paragraph respectively. For my analysis, I create a variable that captures a child's score on each test–these measures range from zero (the child failed to answer any question correctly) to four (she demonstrated the highest level of proficiency). These constitute the main outcome variables of this analysis.

One of the key advantages of using the ASER dataset is its national coverage. The sample size of each survey is large and it encompasses all rural districts in India[7]. Another unique aspect of the ASER data is 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, along with those that are currently enrolled

made under this policy till around late 2004.

[7]See this link for further details on the ASER sampling strategy. In fact, this model has also been adapted for use in some African countries – see UWEZO Surveys.

in school. Additionally, the format of the tests, and the way they are administered and scored have remained uniform across different years and regions, which facilitates spatial and temporal comparisons.

3.2 District Level Household and Facility Survey (DLHS)

I obtain data on DPEP's direct beneficiaries from the DLHS, a household-level survey conducted by the government of India. It is a repeated cross-section that is representative at the district level, and covers 601 districts across 34 Indian states and territories. This survey collates statistics on a wide range of indicators related to household demographic characteristics, maternal and child health, and family planning. I draw upon two rounds of this survey – rounds in 2007-2008 (wave 3) and 2012-2013 (wave 4). I use this data to investigate effect on educational outcomes of direct beneficiaries (years of education, finishing primary school, having ever attended school, and being literate), and to probe mechanisms that could potentially be responsible for DPEP's intergenerational effects – I specifically focus on marriage, fertility and child care related outcomes.

3.3 District Information System for Education (DISE)

To examine school characteristics, I use the District System for Education (DISE), a database that contains information on the physical infrastructure and amenities in government schools, as well as data on teachers, enrolment and other school-level covariates. This source compiles information provided by school headmasters on a yearly basis, and it covers every government primary and upper-primary school, in addition to some private schools. Information provided by schools is verified at the cluster level and subsequently transferred to the district level. At this stage, the data is verified again before being aggregated, digitized and published. I use DISE from the year 2014, and it allows me to test whether the DPEP indeed led to disproportionately more school construction in treatment regions during the years the program was in operation, and to examine treatment-control differences in school quality before, during and after the implementation of DPEP.

3.4 Indian Human Development Survey (IHDS)

The IHDS is a nationally representative panel survey conducted by the University of Maryland in collaboration with the National Council of Applied Economic Research, New Delhi. The first round, IHDS–I, was conducted between November 2004 and October 2005 and covered 41,554 households across 33 states and union territories of India (Desai and Vanneman, 2005). The second wave of the survey (IHDS–II) took place between November 2011 and October 2012, and covered 42,152 households across 1,420 villages and 1,042 urban areas. The sample of the IHDS–II encompassed 83 percent of households from IHDS-I that could be tracked for the follow-up survey. In the current analysis, I primarily use the female module of the survey – this was administered to ever-married women between the ages of 15 and 49, and I use it to probe factors that might have mediated the intergenerational impacts of DPEP. In particular I focus on women's decision making power, their attitudes towards different behaviors by men, and the investment made in children's education.

4 Identification Strategy

In my analysis, I estimate the impacts of DPEP on two different generations – the parent generation (the direct beneficiaries) and the child generation (those who were exposed to the policy indirectly through the parent generation). Below, I describe the empirical frameworks used for these two sets of analyses.

4.1 Direct Beneficiary Effects (First Generation)

To estimate DPEP's effects on direct beneficiaries, I rely on the exogenous variation introduced by the DPEP allocation rule: districts below the cutoff of 39.2 percent female literacy were more likely to receive the policy than districts that were above this cutoff. Leveraging this variation around the RD cutoff (3), I use a Fuzzy Regression Discontinuity (FRD) design to identify the direct effects of the DPEP policy on individuals who were of school-going age at the time of programme implementation. This approach is akin to an instrumental variables (IV) strategy that instruments for DPEP status with a categorical variable for whether an individual resides in a district with lower than national average (39.2 percent) female literacy. It provides the local average treatment effect (LATE) estimate of DPEP, and the equations pertaining to this approach are given below:

$$\textbf{First Stage:} \quad DPEP_{sd} = \alpha_1 BelowAvg_{sd} + f(DFLR_{sd}) + \theta_s + \mu_t + \gamma X_{psdt} + \nu_{psdt} \quad (1)$$

$$\textbf{Second Stage:} \quad Y_{psdt} = \beta_1 DPEP_{sd} + f(DFLR_{sd}) + \theta_s + \mu_t + \delta X_{psdt} + \epsilon_{psdt} \quad (2)$$

where $DPEP_{sd}$ is an indicator variable for whether district d in state s that the individual lived in received the DPEP policy or not[8], and Y_{psdt} represents different outcomes of interest for direct beneficiary (or parent) p, such as years of education, literacy. $BelowAvg_{sd}$ is a dummy variable that takes a value of one if district d in state s has a female literacy rate below the national average (39.2 percent) at the time of program initiation, θ_s represents state fixed effects, μ_t are survey-year fixed effects and X_{psdt} refers to control covariates (caste and religion). Here, $f(DFLR_{sd})$ is a function of the running variable, the District Female Literacy Rate (DFLR). This is measured as the female literacy rate in district d in state s according to the 1991 census (which is the information used to allocate the policy). Here $f(DFLR_{sd})$ takes a quadratic functional form in the main results[9]. The standard errors are clustered at the district level.

The estimation sample consists of individuals who were of primary-school going age at the time of DPEP implementation. In India, primary school consists of grades 1 to 5, and children in primary school tend to be between the ages of 5 and 10 years. However, due to delayed school entry and/or uneven grade progression, even children up to the age of 13 years might remain in primary school (Azam and Saing, 2017). To be conserva-

[8]Ideally, I would assign DPEP status to individuals based on their district of residence when they were of school-going ages (5–14 years of age). But the datasets that I use in this analysis do not have this information – therefore I assign DPEP status based on their current district of residence. Later in the paper, I discuss (and assuage) concerns related to selective migration across districts affecting the impact estimates.

[9]I present several robustness checks where I alter the polynomial functional form. I avoid using higher order polynomials (≥ 3) because they lead to counterintuitive weighting of observations and erratic behavior of the estimator near the boundary point (Gelman and Imbens, 2017). These issues are especially critical in the case of RD estimation since in this analysis I am interested in inference at the boundary point (cutoff value).

tive, I define the sample of direct beneficiaries to include individuals who were between 5 and 14 years of age during the DPEP years. Since a bulk of the school construction under this policy happened between 1994 and 2000, the sample for the direct beneficiary analysis consists of individuals in the primary-school going age range during these years, i.e. those born between the years 1980 and 1995. Later in the text, I check the sensitivity of the results to restricting the sample to individuals who were 5–12 years (born between 1982-1995) during the DPEP years[10].

4.2 Intergenerational Effects (Second Generation)

The second part of the analysis focuses on the intergenerational effects of DPEP. The 2SLS estimation equations in this case are:

$$\textbf{First Stage:} \quad DPEP_{psd} = \alpha_1 BelowAvg_{sd} + f(DFLR_{sd}) + \theta_s + \mu_t + \gamma X_{ipsdt} + \nu_{ipsdt} \tag{3}$$

$$\textbf{Second Stage:} \quad C_{ipsdt} = \beta_1 DPEP_{psd} + f(DFLR_{msd}) + \theta_s + \mu_t + \delta X_{ipsdt} + \epsilon_{ipsdt} \tag{4}$$

where C_{ipsdt} are the different learning outcomes of interest (like math and reading test scores) for child i of parent p residing in district d of state s measured in survey year t, $DPEP_{psd}$ is a categorical variable for whether the parent of the child was a direct beneficiary of DPEP or not, $BelowAvg_{sd}$ is a dummy variable that takes a value of one if the parent of the child resided in a district d in state s where the female literacy rate was below the national average (39.2 percent), θ_s represents state fixed effects, μ_t are survey-year fixed effects and X_{ipsdt} refers to control covariates (gender of the child). The function $f(DFLR_{sd})$ and the clustering of the standard errors are similarly specified.

When probing intergenerational effects, the estimation sample consists of children born to individuals in the direct beneficiary sample (which is described above). Additionally, I restrict the analysis to children who start schooling in 2005 or later, or in other

[10] In appendix A, I discuss other aspects relevant to identification of the estimation sample, including how I address concerns related to migration and changes in district boundary over time.

words, children who themselves would not have been directly affected by DPEP[11]. I discuss the results separately based on the gender of the direct beneficiary – one group of children are those whose mothers benefitted from DPEP, and the others had fathers who benefitted from the policy. As a robustness check, I also replicate the analysis for children with both parents being DPEP beneficiaries.

These estimations are conducted within a bandwidth around the RD cutoff. In the main results, these bandwidths are calculated using two data-driven approaches – the Mean Squared Error (MSE) and the Coverage Error Rate (CER) methods. These methods have been described extensively in Calonico et al., 2014, 2016, 2017, Cattaneo and Vazquez-Bare, 2016, Imbens and Kalyanaraman, 2012. In addition to these data-driven approaches, I also test the robustness of the results to other ways of selecting bandwidth, and to changes in different parameters used in the estimation (like polynomial and kernel function) of the estimation[12].

4.3 Instrument Validity

To be a valid instrument for programme participation, the instrumental variable ($BelowAvg_{sd}$) needs to satisfy two conditions. The inclusion restriction requires that the potentially endogenous independent variable of interest ($DPEP_{sd}$) be correlated with the instrument ($BelowAvg_{sd}$). In other words, the instrument should be a strong predictor of programme participation. Figure 3 plots the probability of a district being part of the treatment group against the 1991 district female literacy rate. The figure clearly illustrates that there is a significant jump or discontinuity in the probability of programme receipt around the programme cutoff of 39.2 percent (the 1991 national average female literacy rate) – districts just below the cutoff were much more likely to be part of the DPEP treatment group as

[11] I restrict my analysis sample to children who would themselves have likely not been directly affected by DPEP. The DPEP programme was phased out starting in 2001-02, but it officially ended around late 2004. Therefore, I select children who are unlikely to have *directly* benefitted from the DPEP. Additionally, in section 5.3, I provide further evidence to show that the observed results are highly unlikely to be the result of direct effects of DPEP on the intergenerational cohort.

[12] Please note that I adopt the continuity-based RD approach in the estimations. In implementing these approaches, I specify the kernel function that is to be used to determine the weight assigned to each observation. In my analysis I use a triangular kernel which puts higher weight on observations close to the RD cutoff and less weight on observations that are further away. Later, I do, however, test whether the results are robust to using an epanechnikov kernel.

compared to districts that were just above the RD threshold[13].

The second condition that is to be met for an instrument to be valid is the exclusion restriction – the instrumental variable (in my case, $BelowAvg_{sd}$) should impact the outcome only through the instrumented variable ($DPEP_{sd}$), and not through other channels[14]. The exclusion restriction is not directly testable, but I argue that it is likely to be satisfied in the setting I examine. To the best of my knowledge, there are no other government programmes (before or after DPEP) that were implemented based on the allocation rule used for the DPEP. Given that there were no discontinuities in the provision of other schemes, it is apriori unlikely that there would be any discontinuities in outcomes around the female literacy cutoff chosen for DPEP. Later in the paper, I conduct several falsification tests to show that there were no discontinuities in variables that should be unaffected by the DPEP. Hence, the instrument that I use is unlikely to be correlated with any other covariates around this cutoff. I discuss this in further detail in the results section.

4.4 RD Validity

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 decide whether or not to be a part of the treatment group. This would lead to non-random assignment to treatment, which would complicate the identification of causal effects. To understand whether such concerns could be playing a role in my setup, I conduct a density test (as described in McCrary, 2008), which helps analyze if there is any *heaping* around the RD cutoff. These results are shown in figure 2. The graph and the associated test statistic (p-value = 0.72) suggest that there is no discontinuity in the forcing variable (district female literacy rate) around the RD cutoff.

[13]Despite their being a significant discontinuity in the probability of treatment assignment, for a variety of implementation-related reasons it is possible that this might not have translated into differences in the actual number of schools constructed across the RD cutoff. In Appendix B, I establish that the DPEP policy was able to achieve its goal of constructing schools in treatment districts by demonstrating the higher number of schools constructed in treatment districts during DPEP years.

[14]In this case this is a conditional exogeneity assumption, where I condition on a smooth function of the running variable on either side of the RD cutoff ($f(DFLR_{msd})$) and a basic set of control covariates.

Could other types of non-random sorting be occurring around the RD cutoff? Theoretically, in this case, such violations could occur in two potential ways – if sub-national governments (at the state or district level) were able to choose their treatment status by manipulating their value of the running variable (DFLR) or if individuals were able to affect their treatment status through systematic migration. The former is unlikely because states (and/or districts) had limited knowledge of the exact decision rule regarding the programme prior to its implementation[15] Additionally, the allocation rule was based on the 1991 census data, which was not only collected well before programme announcement, but was also carried out under the purview of an independent central authority with no state/district oversight. These factors make it next to impossible for sub-national entities to tamper with their value of the running variable.

While individuals could potentially have determined their treatment status through systematic migration across districts, I argue that this was unlikely to have occurred. First of all, permanent migration away from one's location is a rare event – data from a nationally representative household survey (IHDS) suggests that in 2005 more than 93 (96) percent of the households had lived in their current district of residence for more than 10 (5) years. Further, inter-district migration in India mostly takes place for marriage and employment purposes (Topalova, 2007). School choice (especially primary school) was not a major reason for migration in India, especially in rural India, around the time DPEP was implemented in the late 1990s. Any migration to seek enhanced education opportunities, is likely to be confined to the realms of higher education (high school and beyond). Since the DPEP programme mostly constructed primary and upper primary schools, there is low likelihood that systematic migration could have affected the composition of the treatment group in a significant way.

[15] Decisions regarding programme placement were being made by the central government in conjunction with the World Bank and other donors. Also, no other government policies in the past (or since) have been allocated based on the district female literacy rate.

5 Results

5.1 Effect on Parent Generation

I first present the impacts that DPEP had on educational outcomes of direct beneficiaries, individuals who were of school-going age at the time of programme implementation (i.e. those born between 1980 and 1995). Table 1 presents the 2SLS estimates (LATE) for the impact of DPEP on male and female direct beneficiaries separately. I find that the programme had a positive effect on school enrolment, with both males and females experiencing 8-10 percentage point increases. Beneficiaries of both genders also attained more years of education (0.7 - 0.8 years), were more likely to complete primary school (5 to 12 percentage points) and had a higher likelihood of being literate (9 percentage points). The same effects are illustrated in a graphical manner in figure 4 – the upper panel focuses on female beneficiaries and the lower panel on male beneficiaries. These findings corroborate the estimates from table 1, and indicate that people going to school in the immediate aftermath of DPEP implementation attained better educational outcomes[16]. These effects on educational outcomes of direct beneficiaries are similar in magnitude to those found by other studies examining the policy DPEP (Khanna, 2021, Agarwal et al., 2021).

As a robustness check for these results, I also use a difference-in-differences approach to estimating the effect of DPEP on the direct beneficiary group. In these estimations, I leverage two sources of variations – a district's DPEP status, and exposure to the policy depending on the age of the individual during policy years. The results from this analysis are presented in figure 5 – they suggest that cohorts that were 14 years or below at the time of the policy implementation (born 1980 and after) experienced significant gains in years of education as a result of this policy. These results are discussed in greater detail in Appendix C.

To further strengthen the validity of the direct beneficiary effects, I conduct a falsifi-

[16] I also show that this effect on direct beneficiaries is robust to the use of other nationally representative survey datasets. The results presented in figure A1 show that the direct beneficiary effects remain when I use information from the two rounds of Indian Human Development Survey (IHDS).

cation check where I estimate the impact of DPEP on a sub-population that should have been unaffected by this policy – those who were between 18 and 25 years of age at the start of the reform (in 1993). This set of people (born between 1968 and 1975) should be unaffected by DPEP as the policy primarily led to expansion in primary schooling opportunities and participation in primary school at those ages is close to zero. I use the same analytical framework as above. The results are presented in the form of graphs in figure 6, and they are disaggregated by the gender of the beneficiary. They suggest that as expected there were no DPEP impacts on both men and women who had aged-out of benefiting from this policy[17]. This further strengthens our confidence that the results in Table 1 and Figure 4 stem from the DPEP, and reduces concerns that our results might be driven by unobserved differences across DPEP and non-DPEP districts.

5.2 Intergenerational Effects

I now present the results of the effect of DPEP on the learning outcomes of the children of direct DPEP beneficiaries using the empirical strategy outlined in section 4.2. A child could have indirectly been exposed to the consequences of DPEP through either parent (mother or father) or both parents. Since this could have different implications for human capital accumulation, I consider these cases separately.

5.2.1 Mother was DPEP beneficiary

The sample for this part of the analysis consists of children who satisfy the following conditions – (1) have mothers who were born between 1980 and 1995, and were thereby direct DPEP beneficiaries, (2) have fathers who were not direct DPEP beneficiaries, and (3) would themselves have started school after 2005 (after DPEP had been phased-out). Table 2 presents the LATE estimates of the effects of DPEP on test scores for this sample. The test score outcomes are scored in a way such that zero means a failure to provide any correct answers and four indicates complete proficiency on the test. Results in column 1 of table 2 shows that a woman who benefitted from DPEP had a child who scored

[17] This pattern can also be seen in the diff-in-diff results (discussed above). Older cohorts, that is those born in 1978 or earlier, show null effects of the policy (figure 5). These results have been discussed in detail in Appendix C.

0.43 points higher on the reading test than children of comparable non-beneficiary mothers, an effect size that is around 18 and 24 percent of the control group mean standard deviation respectively. These same effects are shown in a graph in the left panel of figure 7. Since a one unit increase on the test implies an increase of one skill level, DPEP's impact on child reading can also be interpreted as an improvement of a little less than half a skill level. Given that an average child in the sample is able to read a word, the identified impact estimate implies that DPEP was able to nudge the child of a typical female program beneficiary towards being able to read a sentence, the next level on the test.

The DPEP impact on math scores of 0.27 points (column 2 in table 2 and left panel of figure 7) moves an average child closer to being able to do a simple addition (score = 3) rather than recognizing a two digit number (score = 2) . This is an improvement of almost 13 (19) percent of the mean (s.d.) math ability in the control group. I also find a positive effect of the programme on English reading ability – the identified RD coefficient is 0.32 points (corresponding graph in figure 7), which is around 15 percent of the mean control group score[18].

How do the effects I identify in this section compare to the intergenerational learning impacts of other policies across developing countries? A study using Indonesian data (Mazumder et al., 2021) finds that mother's exposure to a school construction programme leads to higher test scores for children, especially daughters. The authors find effect sizes in the order of 0.07 S.D. (σ) to 0.1 σ – these are comparable but smaller than the effects that I find evidence for in the current study. The effect sizes I identify are, however, smaller but similar to effects uncovered by other investigations of school construction programmes[19]: 0.65 σ (Burde and Linden, 2013), around 0.5 σ (Banerjee et al., 2007), 0.2 σ (Glewwe et al., 2009), around 0.2 σ (Duflo et al., 2012), around 0.3 σ (Muralidharan et al., 2019) and 0.4-2.2 σ (Kazianga et al., 2013).

[18]Note that I do not claim that the observed effects are solely driven by parental education, especially as Khanna [2021] demonstrates that DPEP had large general equilibrium effects, especially on wages of direct beneficiaries.

[19]These are largely effects on direct beneficiaries of the respective policies, and not the intergenerational effects that I identify. Also, σ here represents standard deviation

5.2.2 Father was DPEP beneficiary

Akin to the analysis above, I examine outcomes for children whose fathers only directly benefitted from DPEP. The results in Table 3 suggest that the point estimates of the impact on this sample are much smaller in magnitude compared to the mother beneficiary group, and are not statistically significant. The null results are also visible in the graphs – the right panel of figure 7 also shows that there is no discontinuity in the test score outcome around the RD cutoff for children of father beneficiaries. Recall, that the fathers of these children did benefit from the DPEP – for example, by attaining more education than their counterparts in non-DPEP districts (Table 1 and the right panel of figure 5). The null results for the children of male direct beneficiaries are also unlikely to stem from small sample sizes within the bandwidth used to generate these impact estimates – the effective number of observations are more than 26,000 for the outcomes examined (except English score, in which case it is roughly 7,200). Rather, it seems likely that fathers, unlike mothers, were unable to use their enhanced schooling to improve the outcomes of their children. This is in line with the findings of Mazumder et al. [2021] in the Indonesian context.[20]

5.2.3 Gender heterogenous effects

In this section, I examine whether the transmission of human capital across generations varies by the gender of the child. Figure 8 conveys the following trends: first, both daughters and sons gain when mothers benefitted from DPEP. Second, the effects transmitted to children through father beneficiaries are small and statistically insignificant, irrespective of the gender of the child. Third, the differences between the effects of mothers and fathers on their children (of either gender) are substantial, but these differences themselves are not statistically significant (the confidence intervals overlap). This is consistent with the findings of Burde and Linden [2013] (in Afghanistan) and Mazumder et al. [2021] (in Indonesia), both of which study the impact of schooling expansion policies[21].

[20] Table A3 presents results for the sample of children who had both parents benefiting from the DPEP program. The results are qualitatively similar to the results for the children with only mothers exposed to DPEP school construction (Table 2).

[21] Broadly speaking, these results are in line with other studies that find the role of mother to be vital in the human capital formation of children (Desai and Alva, 1998, Currie and Hyson, 1999, Currie and

5.3 Establishing DPEP had no *direct* effect on intergenerational sample

To credibly establish that the second generation effects that I observe are only due to parental DPEP exposure, I need to rule out that DPEP had any direct effects on the children of the main program beneficiaries. In this section, I explore some of these other channels through which DPEP could have directly affected the child sample I examine in this analysis.

Although the DPEP policy had been phased out by the time that the second-generation children entered school, the infrastructural changes brought about by DPEP would still be in place. Any resulting systematic differences in school infrastructure between DPEP and non-DPEP regions could have led to the enhanced learning outcomes that I observe for the second-generation sample. I contend, however, that DPEP-induced infrastructure changes are unlikely to differentially impact the children of the direct program beneficiaries, and below I discuss some of the reasons why.

Using school-level information from more than a million schools across India, I first establish that there were no significant quality or quantity differences between schools in DPEP and non-DPEP districts in the post-programme years (years 2006 and beyond)[22]. Figures A2 and A3 depict the RD estimate of the *effect* of DPEP on two measures of school quantity – the rate of annual school construction and the total number of schools, respectively. These estimates are presented for the years 2006 to 2012, the period when the second-generation beneficiaries were attending school. They show that there were no significant differences in any measure of school quantity across the RD cutoff in post-DPEP period. School quality can be represented by school infrastructure (number of classrooms, and presence of a toilet/girls toilet/electricity), teacher characteristics (number of male/female teachers, teachers with a graduate degree, professional qualification), school oversight (distance to block/cluster office, number of inspection visits made by

Madrian, 1999, Persico et al., 2004, Case et al., 2002, Behrman and Rosenzweig, 2004, Case et al., 2005, King et al., 2007, Güneş, 2015, Vollmer et al., 2016, Alderman and Headey, 2017).

[22] This information comes from the District Information System for Education (DISE) dataset. Note, that this dataset contains detailed information on government, government-aided and a few private schools in nearly two-thirds of the districts across the country.

block/cluster officials) and receipt of funding/grants. I argue that put together, these outcomes provide a comprehensive outlook of schooling quality, and I examine these in table A2. The results suggest that there were no statistically significant differences between DPEP and non-DPEP regions along any of these characteristics once the program had come to its end (see Appendix B for a more detailed discussion of these results).

Despite the results discussed above, one could argue that DPEP could have led to improvements in unobserved aspects of school quality in districts targeted by the program and that these factors were responsible for the learning gains observed for the children of DPEP direct beneficiaries[23]. If DPEP did have such far reaching changes, then one would expect that on average, children in DPEP districts would have better learning outcomes than their counterparts in non-DPEP regions in the post-programme period. If so, then this would be true even for children of DPEP non-beneficiaries – individuals who were between 18 and 25 years of age at the time of DPEP initiation in 1993 (those born between 1968 and 1975), and hence would have aged-out of primary school and thus failed to come within the purview of this reform[24]. Specifically, I conduct a check where I examine whether there were differences in learning outcomes of children of these non-beneficiaries. If this group shows any statistically significant positive DPEP impacts, then that might suggest that there were channels other than parental DPEP exposure that could potentially drive the observed results. Figure 9 depicts the impacts on this sample of children. The panel on the left (right) is for children who satisfied two conditions – (1) whose mothers (fathers) were 18-25 years of age in 1993 (born between 1968 and 1975) and (2) they started school after 2005 (end of DPEP). For both samples, I fail to find any evidence of statistically significant differences across DPEP and non-DPEP districts – this provides additional credence to the argument that the intergenerational learning effects that I observe in the main analysis are driven by parents' exposure to DPEP and not by other observable (or unobservable) factors.

[23] Such a situation is fairly unlikely since existing evidence from developing countries suggests that there is limited association between school quality and the learning outcomes of children (Glewwe et al., 2011).

[24] Earlier in the analysis I have shown that this group experienced none of the benefits of DPEP (figure 6).

5.4 Robustness Checks

I check the sensitivity of the findings to changes in different parts of the empirical strategy used in tables 2 and 3, such as variations in the bandwidth selection, the polynomial function used to account for the running variable and the kernel function used to assign weights to the observations around the cutoff.

Approach to RD estimation: Recall, that I use the CER approach to bandwidth estimation in my main results (like in table 2). In this check, I instead use the MSE approach to optimal bandwidth estimation. I separately estimate RD treatment effects using a linear and a quadratic polynomial of the running variable (district female literacy rate). These results are presented in figure A4 – the left panel shows the coefficient estimates for children whose mothers were DPEP beneficiaries and the right panel shows estimates for the children whose fathers were beneficiaries. I find the same pattern of results as in the main results – mothers' DPEP status has a statistically significant impact on their children's math and reading scores, whereas beneficiary fathers have a lower positive but statistically insignificant effect on their children.

Choice of bandwidth: In the main results, the bandwidths are selected using data-driven procedures. In the next check, I use an *ad-hoc* approach to defining the estimation neighborhood around the cutoff. Specifically, I choose bandwidth sizes starting with three percent of the district female literacy rate, and gradually moving up by an increment of one until I arrive at 12 percent of the literacy rate. The results from this exercise are presented in figures A5 (reading scores), A6 (math scores) and A7 (English scores). The upper panel of each figure presents point estimates for children of female beneficiaries – the coefficients are always positive and statistically significant. The lower panels are for the children of male beneficiaries – estimates are close to zero (largely positive) and statistically insignificant. This shows that the results are robust to the use of different bandwidths around the cutoff.

Choice of kernel function: While, I use a triangular kernel for the main results, I reexamine the results using an epanechnikov kernel. Figure A8 plots coefficients when an epanechnikov kernel is paired with both the data-driven bandwidth procedures (MSE and CER), and with linear or quadratic polynomial functions of the forcing variable. The point estimates do change marginally, but for the most part retain their sign and significance.

Definition of treatment group: For the next test, I alter the way I define the cohort of individuals who would have directly benefitted from DPEP. In the main set of results, I define children between 5 and 14 years during the DPEP years to be the direct beneficiary group (born between 1980 and 1995) of the school construction program. In India, children typically start school at age five and are required to complete five years of primary schooling, but because of delayed school entry and grade repetition, it is not uncommon for children to stay in primary school till the age of 13 (Azam and Saing, 2017). This reason motivates my use of 14 years to define the upper limit of the cohort of children potentially exposed to DPEP, a program that mostly constructed primary (and upper-primary) schools . If delayed school entry and grade repetition were not substantial, I might however be including a non-trivial sample of non-treated individuals in my sample of direct beneficiaries, thus underestimating the true effect of the policy[25]. Therefore, I re-estimate the results with the lower age cutoff of 12 years to define the cohort that might have been impacted by the programme. In other words, I assume that the direct beneficiary group consists of individual born between 1982 and 1995[26]. The results from this exercise are presented in figure A9 – the RD coefficients retain both their sign and statistical significance, implying that the main study findings persist when using a different (and potentially more precise) definition of the treatment group.

[25]The diff-in-diff results in figure 5 also suggest this – cohorts born in 1981 and 1982 exhibit modest education effects of exposure to DPEP.

[26]This lower age cutoff is especially relevant for girls since they tend to drop out of schools at younger ages than boys due to a variety of reasons, such as child marriage and onset of menarche (Fentiman et al., 1999, Roberts et al., 2002, Burrows 1 and Johnson, 2005, Sommer, 2010, Adukia, 2017).

5.4.1 Accounting for delays in DPEP implementation

DPEP was announced in 1993-94, and so in the main analysis I consider everyone in treatment regions who was of primary school–going age (5 to 14 years) at this point in time to be a direct beneficiary of DPEP, an approach similar to the one used by Khanna, 2021 who also evaluates the same policy (though on different outcomes). It is, however, worth recognizing that given the inherent challenges and the wide geographical spread of DPEP, the program began at different times in districts across the country[27]. Not accounting for these differences might thus lead to bias in treatment effects.

In this section, I attempt to account for these differences by using multiple methods to assign DPEP start date to each district. First, I use government documents to identify the year of DPEP initiation in each treatment district (subsequently referred to as the *archival start date*), and accordingly identify the direct beneficiary sample. This process requires triangulating information on programme expenditures, field reports on programme implementation and progress reports created at the state or central government level. For example, I consider DPEP to have begun in a district if money for school construction had been sent to the district (indicated by documents of the central government) and it had been spent by the state/district authorities (noted in state expenditure reports) and if school construction had begun (specified in annual progress reports prepared by state and central government). I re-estimate the intergenerational effects of DPEP using the main RD setup with one key difference – the sample used in the analysis now consists of primary school-going age children (5-14 year olds) in each district based on the years of operation of DPEP in that particular district using the *Archival start dates*. The results based on this method are presented in figure A10 – they are consistent with the main set of results[28].

[27]Unexpected interruptions came in the years 1996, 1998 and 1999 when general elections, which are typically held once in five years, were conducted. School personnel had to be on election duty. There were also some instances of delays in funding flows, especially in the early years of the policy.

[28]In this case, to form the comparison group I need to assign programme years to non-DPEP regions. I do this in multiple ways. In table A10, I use the same comparison group as the main analysis. I use another method to assign a likely DPEP start year for the control districts. Under this method, for each state I take the average starting year of DPEP in treatment districts within that state, and consider this year to be the start date for all non-DPEP districts within the state. This method accounts for the state (or regional) differences in implementation patterns across different districts. To show the robustness of my results, I

The *archival data* analysis I describe above involves manual parsing of hundreds of government documents, and has scope for human error. Therefore, I use another more automated technique to infer the start date of the policy across the country. Specifically, I use information from an independent government database called the District Information System for Education (DISE) to calculate the annual rate of growth of schools over time in districts. Since DPEP was meant to boost school construction in treatment districts, one would expect that after the implementation of this policy, there would be a break in the district-specific long term rate of growth of schools. I assign new DPEP start dates (referred to as *DISE start years*) to districts based on the year in which the treatment districts experienced the highest rate of growth in schools[29]. The main results persist when using this alternative approach for defining DPEP treatment (see figure A12).

As a caveat, I note that the start years based on the *Archival data* and the *DISE* information may be endogenous – there might be systematic differences between early and late adopting districts that might be correlated with the outcomes of interest. Therefore, the robustness checks in this section are solely meant to verify the sensitivity of the main results, and the point estimates identified from these robustness checks should be interpreted with caution.

6 Mechanisms

There are multiple potential pathways through which a school construction programme (like DPEP) could shape intergenerational learning outcomes. While I show that individ-

replicate the main results using this alternative definition of identifying the control group (figure A11). In another approach, I use the average of all DPEP district start times as the *start* time for non-DPEP districts. The results based on this exercise also show similar patterns to that of the main results (they are available from the author on request).

[29]Figure A13 illustrates how I assign program start dates using DISE. Each graph plots the annual rate of growth of schools in a given district. The dashed (red) line shows the year when the DPEP policy was announced (1993), the bold (black) line represents the year of implementation inferred from the archival documents (the Archival start year) and the peak of the rate of growth distribution is an alternative start date (the DISE start year) of the policy. These graphs illustrate that treatment districts experienced a break in the long-term trend of school construction after the DPEP programme was announced. Additionally, in figure A14 I plot the same for some non-DPEP districts – as expected, I do not observe any such *breaks* in the annual rate of school construction in these districts.

uals directly exposed to DPEP were able to increase their school attainment (Table 1), I don't know what it was about this schooling that enabled them to positively impact their children's learning outcomes in subsequent years. I now examine several potential pathways that could have been responsible for the observed intergenerational effects. Given that female DPEP beneficiaries appear to be most able to use their education to shape their children's lives, here I focus on DPEP's female beneficiaries and children who had mothers, but not fathers, benefiting from the program.

6.1 Educational Investments

It is possible that parents with high levels of education invest more in their children's education than less educated parents. I use a nationally representative household survey dataset (IHDS) to examine whether DPEP's intergenerational effects on education could have been transmitted through such a channel. One way that parents can enhance investments in child schooling is by enrolling children in potentially higher quality schools, which in the Indian context often tend to be private schools[30]. I estimate whether DPEP exposure had an impact on parents' choice between private and government schools and find that there was no statistically significant impact on private school enrolment (Table 4). Investment in child schooling could also take the form of education expenditures (for example - on books, tuition etc.). Results in table 4 suggest that while DPEP beneficiaries spend no more on school fees, they allocate more resources towards the purchase of books and uniforms (40 percent relative to mean) and private tuition services (30 percent relative to mean) for their children[31]. In addition, the children of programme beneficiaries spend around two more hours doing homework per week) than comparable children of non-beneficiaries, an improvement that might be emerging due to increased supervision by educated mothers. My findings are similar to the results found by Andrabi et al. [2012] in Pakistan, which is culturally similar to the context I examine in the current analysis,

[30] Evidence from India shows that private school attendance in India leads to large improvements on English test scores and a moderate impact on mathematics and vernacular test scores (Singh, 2015). This is despite the fact that private schools pay teachers lesser and spend less per pupil than government schools (Desai et al., 2009, Kingdon, 2007, Muralidharan and Kremer, 2006).

[31] This is in line with the findings of Nordman et al., who show that fall in education expenditure that occurs as a result of agricultural productivity shocks (and potential income loss) are mitigated by increased mothers education.

where they find that children of mothers with some education spend more time on educational activities outside of school hours. Overall, therefore, there is some evidence that DPEP's intergenerational impacts might be mediated by higher parental investments in children's education.

6.2 Child Care Investments

A child's human capital is significantly impacted by in-utero and early life conditions, or what is known as the first 1000 days of life (Almond and Currie, 2011 and Currie and Vogl, 2013 provide comprehensive reviews of this literature). Did DPEP exposure lead to higher usage of Ante Natal Care (ANC) and Post Natal Care (PNC) and thereby improve early life conditions for children? Such services could lead to better health outcomes for children as well as for the women themselves (Paudel et al., 2014, Onasoga et al., 2012, Simkhada et al., 2008, Kerber et al., 2007). In Table 5, I find that the women DPEP beneficiaries were more likely to make at least one ANC visit (7 p.p.), make more ANC visits in total (0.22 visits), deliver in a formal health facility (6 p.p.) and receive post-natal care (6 p.p.). Based on these findings, it is clear that women who were impacted by DPEP were more likely to receive care during and after their pregnancy, which arguably could have led to the enhanced child human capital later in life.

6.3 Health of Direct Beneficiaries

It is well established that maternal health (and health behaviors) is a key determinant of child health and well-being (Ahlburg, 1998, Coneus and Spiess, 2012, Bhalotra and Rawlings, 2013, Yan, 2015). Therefore, it is plausible that DPEP's positive effect on female education had a knock-on effect on female health, and the latter led to higher well- being of children. I test this hypothesis using data from the National Family Health Survey (2015-16), which is the latest nationally representative dataset on health outcomes at the time of analysis. I find that DPEP female beneficiaries indeed have better health as adults as measured by height, BMI and hemoglobin (figure A15). Additionally, I find that beneficiaries are less likely to be anemic or obese. This is in line with a large literature that finds evidence of a positive impact of women's education on their own health in a variety of contexts (Grossman, 2015, Grépin and Bharadwaj, 2015, Agüero and Bharadwaj, 2014, Lundborg, 2013, Amin et al., 2013, Silles, 2009, Currie and Moretti, 2003). In addition, I also find that DPEP had a beneficial impact on female contraceptive usage (table 5), an outcome that might reduce unwanted fertility and thus foster higher human capital of children (Kugler and Kumar, 2017). This finding too resonates with those from other studies (Johnston et al., 2015, Andalón et al., 2014).

6.4 Marriage Outcomes & Bargaining Power

Studies show that women who stay in school longer tend to marry at a higher age, and consequently experience improved outcomes in adulthood, such as enhanced bargaining power, autonomy and labor market outcomes (Lundberg and Pollak, 1993, Field and Ambrus, 2008, Duflo, 2012, Samarakoon and Parinduri, 2015, Crandall et al., 2016, Sunder, 2019, Yount et al., 2018). As the next set of potential mechanisms, I test whether DPEP impacted women's marriage outcomes and bargaining power (using IHDS data). The results in Table 5 show that women beneficiaries married about half a year later and had their first birth 0.33 years later than women in the control group. I also examine the social status of female beneficiaries in the households they married into and find that the DPEP women are 8-14 p.p. more likely to participate in decisions related to their children and household meals (Table 4). These women are also less likely to report that physical violence by husbands against wives is justified (5-9 p.p.). Since women who benefitted from DPEP appear to have higher bargaining power and autonomy within their households, they might be able to shape their children's outcomes more effectively than non-beneficiaries (Yoong, 2012, Bono et al., 2016)[32].

[32] A bulk of the women direct beneficiaries in the sample are still in the age range where they might have more kids (their reported fertility history at the time of the survey may not represent the total number of children that they will have in their lives). Therefore, looking at the impact on number of siblings of the children in our sample would likely present an incomplete picture. Having noted this caveat, I use the main empirical specification laid out in the analysis with the DLHS-3 and DLHS-4 data to examine differences in number of siblings that the second generation children have – the impact estimates of DPEP on total number of siblings are -0.04 (p-value = 0.31) and -0.07 (p-value = 0.26) respectively.

6.5 Relative contribution of mechanisms

I make back-of-the-envelope calculations of the relative contribution that different channels have on the observed effect on reading, math and English tests scores. To do this, I combine the following – (1) the LATE effect of DPEP on the outcome of interest, (2) the LATE effect of the DPEP on each mechanism, and (3) the association between the mechanisms and the outcome of interest. To obtain the relative contribution of each mechanism (provided in table A5), I divide the effect of DPEP on test scores (column 4) by the product of columns (2) and (3). The results suggest that marriage age and age at first birth contributes relatively little to the effect on test scores – each of them contributes around 1-3 percent of the total observed effect. This implies that marriage and fertility outcomes are unlikely to be the main drivers of the observed effect. Instead, I find that women's decision making power (8-15 percent) and education-related expenditure on children (21-25 percent) might have much larger roles to play in mediating the observed effect of DPEP on intergenerational test scores[33].

7 Conclusion

In this paper, I use exogenous variation introduced by the allocation rule of a national-level school construction programme in India to conduct a fuzzy RD analysis to estimate how the program impacted parental schooling and subsequently shaped intergenerational learning outcomes. I find that both male and female individuals exposed to the program in treatment districts were more likely to be literate and attain more education than comparable individuals in districts that did not receive the program. Further, I find that children of female beneficiaries performed better on vernacular reading, math and English test scores than the children of female non-beneficiaries. In contrast, male beneficiaries were unable to transfer the educational gains they experienced to their children. I conduct multiple robustness checks to establish the internal validity of my results. I also

[33]The association between different channels and the outcomes of interest (test scores) is purely associational in nature. This is because the research design used in this study is not adequate to establish causality in these relationships (and can only be used to explore causal effect of DPEP on different outcomes). I also note that these associations provide a reduced form effect of the channel, and hence may capture other factors that are associated with these mechanisms.

validate the results through several placebo tests – for example, I show that individuals who were too old to benefit from DPEP show no effects of the programme and their children also demonstrate no learning gains.

In exploring the potential mechanisms through which the intergenerational impacts of the school construction could have been mediated, I find that women who were able to enhance their educational attainment through DPEP have better health (height, BMI and hemoglobin) and practice superior health behaviors in terms of contraceptive usage, pre-natal care and post-natal care as compared to non-beneficiary women. I also find DPEP's female beneficiaries marry later, and have higher bargaining power in their marital households. All these factors are likely to enable women to allocate greater resources towards their children's welfare. In fact, I find that the children of these women benefit from higher spending on uniforms/books and outside-school private lessons.

In discussing the potential generalizability of the study findings, I note two key points - first, that although all impact estimates are based on comparing individuals in districts close to the program cut-off, the study sample within these bandwidths consists of individuals from different parts of India, which makes the results nationally relevant. Second, there are many countries in Africa (like Ethiopia and Ivory Coast) and South Asia (Pakistan and Afghanistan) that have female literacy levels close to or lower than the RD cutoff (39.2 percent) in this study. Therefore, even though I am able to identify the Local Average Treatment Effects of DPEP, arguably, I am able to do so at a point in the female literacy distribution which approximates those prevailing in many developing countries. As a result, the results from this analysis can possibly inform education policies in these parts of the world, and are thereby externally relevant.

Cognitive development and learning in childhood have an important bearing on later life outcomes. While the bulk of the literature has focused on school based reforms to improve learning outcomes, there have been some demand-side interventions that have been found to be effective – provision of accurate information on returns to schooling, conditional cash transfers and merit scholarships (Glewwe and Muralidharan, 2016 provides a review of the associated literature). Through the current analysis, I demonstrate that parents, particularly mothers, are able to use their education to shape their children's ability to learn. As these (and other findings in Andrabi et al., 2012, Banerji et al., 2017) suggest, improving improving educational opportunities for women could go a long way in boosting child performance on cognitive tests. Therefore, there is a need for policies that target parents to potentially increase educational investment in their children (Andrabi et al., 2015, Bergman, 2015). These reforms should complement, and not substitute, the school-based reforms aimed at improving child learning.

References

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

Madhuri Agarwal, Vikram Bahure, and Sayli Javadekar. Marrying young: The surprising effect of education. *Working Paper*, 2021.

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.

Jorge M Agüero and Maithili Ramachandran. The intergenerational transmission of schooling among the education-rationed. *Journal of Human Resources*, 55(2):504–538, 2020.

Dennis Ahlburg. Intergenerational transmission of health. *The American Economic Review*, 88(2):265–270, 1998.

Richard Akresh, Daniel Halim, and Marieke Kleemans. Long-term and intergenerational effects of education: Evidence from school construction in Indonesia. Technical report, National Bureau of Economic Research, 2018.

Harold Alderman and Derek D Headey. How important is parental education for child nutrition? *World Development*, 94:448–464, 2017.

Harold Alderman, Jooseop Kim, and Peter F Orazem. Design, evaluation, and sustainability of private schools for the poor: the Pakistan urban and rural fellowship school experiments. *Economics of Education Review*, 22(3):265–274, 2003.

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.

Tahir Andrabi, Jishnu Das, and Asim Ijaz Khwaja. What did you do all day? maternal education and child outcomes. *Journal of Human Resources*, 47(4):873–912, 2012.

Tahir Andrabi, Jishnu Das, and Asim Ijaz Khwaja. *Report cards: The impact of providing school and child test scores on educational markets*. The World Bank, 2015.

Mehtabul Azam and Chan Hang Saing. Assessing the impact of district primary education program in India. *Review of Development Economics*, 21(4):1113–1131, 2017.

Sarah Baird, Craig McIntosh, and Berk Özler. Cash or condition? Evidence from a cash transfer experiment. *The Quarterly Journal of Economics*, 126(4):1709–1753, 2011.

Abhijit V Banerjee, Shawn Cole, Esther Duflo, and Leigh Linden. Remedying education: Evidence from two randomized experiments in India. *The Quarterly Journal of Economics*, 122(3):1235–1264, 2007.

Rukmini Banerji, James Berry, and Marc Shotland. The impact of maternal literacy and participation programs: Evidence from a randomized evaluation in India. *American Economic Journal: Applied Economics*, 9(4):303–37, 2017.

World Bank. Project performance assessment report - India. 40160-IN, 2007.

Felipe Barrera-Osorio, Marianne Bertrand, Leigh L Linden, and Francisco Perez-Calle. Improving the design of conditional transfer programs: Evidence from a randomized education experiment in Colombia. *American Economic Journal: Applied Economics*, 3(2): 167–95, 2011.

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.

Jere R Behrman and Mark R Rosenzweig. Does increasing women's schooling raise the schooling of the next generation? *American Economic Review*, 92(1):323–334, 2002.

Jere R Behrman and Mark R Rosenzweig. Returns to birthweight. *Review of Economics and Statistics*, 86(2):586–601, 2004.

Jere R Behrman, David Ross, and Richard Sabot. Improving quality versus increasing the quantity of schooling: Estimates of rates of return from rural Pakistan. *Journal of Development Economics*, 85(1-2):94–104, 2008.

Jere R Behrman, Susan W Parker, and Petra E Todd. Schooling impacts of conditional cash transfers on young children: Evidence from Mexico. *Economic development and cultural change*, 57(3):439–477, 2009.

Julia Andrea Behrman. Does schooling affect women's desired fertility? Evidence from malawi, Uganda, and ethiopia. *Demography*, 52(3):787–809, 2015.

Peter Bergman. Parent-child information frictions and human capital investment: Evidence from a field experiment. 2015.

Sonia Bhalotra and Samantha Rawlings. Gradients of the intergenerational transmission of health in developing countries. *Review of Economics and Statistics*, 95(02):660–672, 2013.

Paul Bingley, Kaare Christensen, and Ian Walker. Twin-based estimates of the returns to education: Evidence from the population of Danish twins. *Working Paper*, 2005.

Paul Bingley, Kaare Christensen, and Vibeke Myrup Jensen. Parental schooling and child development: Learning from twin parents. 2009.

Anders Björklund and Kjell G Salvanes. Education and family background: Mechanisms and policies. In *Handbook of the Economics of Education*, volume 3, pages 201–247. Elsevier, 2011.

Sandra E Black, Paul J Devereux, and Kjell G Salvanes. Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *American Economic Review*, 95(1):437–449, 2005.

Moussa P Blimpo. Team incentives for education in developing countries: A randomized field experiment in benin. *American Economic Journal: Applied Economics*, 6(4):90–109, 2014.

Emilia Del Bono, Marco Francesconi, Yvonne Kelly, and Amanda Sacker. Early maternal time investment and early child outcomes. *The Economic Journal*, 126(596):F96–F135, 2016.

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.

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.

Pedro Carneiro, Costas Meghir, and Matthias Parey. Maternal education, home environments, and the development of children and adolescents. *Journal of the European Economic Association*, 11(suppl_1):123–160, 2013.

Anne Case, Darren Lubotsky, and Christina Paxson. Economic status and health in childhood: The origins of the gradient. *American Economic Review*, 92(5):1308–1334, 2002.

Anne Case, Angela Fertig, and Christina Paxson. The lasting impact of childhood health and circumstance. *Journal of Health Economics*, 24(2):365–389, 2005.

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

Raj Chetty, John N Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. How does your kindergarten classroom affect your earnings? Evidence from project star. *The Quarterly Journal of Economics*, 126(4):1593–1660, 2011.

Arnaud Chevalier, Colm Harmon, Vincent O'Sullivan, and Ian Walker. The impact of parental income and education on the schooling of their children. *IZA Journal of Labor Economics*, 2(1):8, 2013.

Shin-Yi Chou, Jin-Tan Liu, Michael Grossman, and Ted Joyce. Parental education and child health: Evidence from a natural experiment in Taiwan. *American Economic Journal: Applied Economics*, 2(1):33–61, 2010.

Katja Coneus and C Katharina Spiess. The intergenerational transmission of health in early childhood—evidence from the german socio-economic panel study. *Economics & Human Biology*, 10(1):89–97, 2012.

AliceAnn Crandall, Kristin VanderEnde, Yuk Fai Cheong, Sylvie Dodell, and Kathryn M Yount. Women's age at first marriage and postmarital agency in Egypt. *Social Science Research*, 57:148–160, 2016.

Janet Currie and Rosemary Hyson. Is the impact of health shocks cushioned by socioeconomic status? the case of low birthweight. *American Economic Review*, 89(2):245–250, 1999.

Janet Currie and Brigitte C Madrian. Health, health insurance and the labor market. *Handbook of Labor Eonomics*, 3:3309–3416, 1999.

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. *Annual Review of Economics*, 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.

Sonalde Desai and Soumya Alva. Maternal education and child health: Is there a strong causal relationship? *Demography*, 35(1):71–81, 1998.

Sonalde Desai and Reeve Vanneman. National council of applied economic research, new delhi. 2005. *India Human Development Survey (IHDS)*, 2005.

Sonalde Desai, Amaresh Dubey, Reeve Vanneman, Rukmini Banerji, et al. Private schooling in India: A new educational landscape. In *India Policy Forum*, volume 5, pages 1–38. National Council of Applied Economic Research, 2009.

Esther Duflo. Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review*, 91 (4):795–813, 2001.

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.

Esther Duflo. Women empowerment and economic development. *Journal of EconomIc Literature*, 50(4):1051–79, 2012.

Esther Duflo, Rema Hanna, and Stephen P Ryan. Incentives work: Getting teachers to come to school. *American Economic Review*, 102(4):1241–78, 2012.

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.

Daniel O Gilligan, Naureen Karachiwalla, Ibrahim Kasirye, Adrienne M Lucas, and Derek Neal. Educator incentives and educational triage in rural primary schools. *Journal of Human Resources*, 57(1):79–111, 2022.

Animesh Giri and Vinish Shrestha. The effect of school construction on educational outcomes among females: Evidence from Nepal. *Development Journal of the South*, 2(1):1, 2017.

Paul Glewwe. The relevance of standard estimates of rates of return to schooling for education policy: A critical assessment. *Journal of Development Economics*, 51(2):267–290, 1996.

Paul Glewwe and Karthik Muralidharan. Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications. In *Handbook of the Economics of Education*, volume 5, pages 653–743. Elsevier, 2016.

Paul Glewwe, Michael Kremer, and Sylvie Moulin. Many children left behind? textbooks and test scores in Kenya. *American Economic Journal: Applied Economics*, 1(1):112–35, 2009.

Paul W Glewwe, Eric A Hanushek, Sarah D Humpage, and Renato Ravina. School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010. Technical report, National Bureau of Economic Research, 2011.

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.

Pınar Mine Güneş. The role of maternal education in child health: Evidence from a compulsory schooling law. *Economics of Education Review*, 47:1–16, 2015.

Sudhanshu Handa. Raising primary school enrolment in developing countries: The relative importance of supply and demand. *Journal of Development Economics*, 69(1):103–128, 2002.

Eric A Hanushek and Ludger Woessmann. The role of cognitive skills in economic development. *Journal of Economic Literature*, 46(3):607–68, 2008.

Eric A Hanushek and Ludger Woessmann. Do better schools lead to more growth? cognitive skills, economic outcomes, and causation. *Journal of Economic Growth*, 17(4):267–321, 2012.

James J Heckman. Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782):1900–1902, 2006.

James J Heckman. Schools, skills, and synapses. *Economic Inquiry*, 46(3):289–324, 2008.

Helena Holmlund, Mikael Lindahl, and Erik Plug. The causal effect of parents' schooling on children's schooling: A comparison of estimation methods. *Journal of EconomIc Literature*, 49(3):615–51, 2011.

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.

Robert Jensen. The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, 125(2):515–548, 2010.

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.

Dean Jolliffe. Skills, schooling, and household income in Ghana. *The World Bank Economic Review*, 12(1):81–104, 1998.

Heidi Kaila, David Sahn, and Naveen Sunder. What you learned by second grade matters: A comparative study on human capital formation in Madagascar and Senegal. *Working Paper*, 2020.

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

Kate J Kerber, Joseph E de Graft-Johnson, Zulfiqar A Bhutta, Pius Okong, Ann Starrs, and Joy E Lawn. Continuum of care for maternal, newborn, and child health: from slogan to service delivery. *The Lancet*, 370(9595):1358–1369, 2007.

Gaurav Khanna. Large-scale education reform in general equilibrium: Regression discontinuity evidence from India. *Working Paper, University of California - San Diego*, 2021.

Elizabeth M King, Stephan Klasen, and Maria Porter. *Women and development*. Copenhagen Consensus Center, 2007.

Geeta Gandhi Kingdon. The progress of school education in India. *Oxford Review of Economic Policy*, 23(2):168–195, 2007.

Michael Kremer, Conner Brannen, and Rachel Glennerster. The challenge of education and learning in the developing world. *Science*, 340(6130):297–300, 2013.

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.

Krishna Kumar, Manisha Priyam, and Sadhna Saxena. Looking beyond the smokescreen: Dpep and primary education in India. *Economic and Political Weekly*, pages 560–568, 2001.

Tao Li, Li Han, Linxiu Zhang, and Scott Rozelle. Encouraging classroom peer interactions: Evidence from Chinese migrant schools. *Journal of Public Economics*, 111:29–45, 2014.

Dajun Lin, Randall Lutter, and C Ruhm. What are the effects of cognitive performance on labor market outcomes? *University of Virginia, July. http://batten. virginia. ed u/sites/default/files/research/attachments/The% 20Effects% 20of% 20Cognitive*, 20, 2016.

Prashant Loyalka, Chengfang Liu, Yingquan Song, Hongmei Yi, Xiaoting Huang, Jianguo Wei, Linxiu Zhang, Yaojiang Shi, James Chu, and Scott Rozelle. Can information and counseling help students from poor rural areas go to high school? Evidence from China. *Journal of Comparative Economics*, 41(4):1012–1025, 2013.

Shelly Lundberg and Robert A Pollak. Separate spheres bargaining and the marriage market. *Journal of Political Economy*, 101(6):988–1010, 1993.

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

Francesca Marchetta and David E Sahn. The role of education and family background in marriage, childbearing, and labor market participation in Senegal. *Economic Development and Cultural Change*, 64(2):369–403, 2016.

Bhashkar Mazumder, Maria Fernanda Rosales-Rueda, and Margaret Triyana. Social interventions, health and wellbeing: The long-term and intergenerational effects of a school construction program. *Journal of Human Resources*, pages 0720–11059R1, 2021.

Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714, 2008.

Peter G Moll. Primary schooling, cognitive skills and wages in south africa. *Economica*, 65(258):263–284, 1998.

Karthik Muralidharan. Priorities for primary education policy in India's 12th five-year plan. In *India Policy Forum*, volume 9, pages 1–61. National Council of Applied Economic Research, 2013.

Karthik Muralidharan and Michael Kremer. Public and private schools in rural India. *Harvard University, Department of Economics, Cambridge, MA*, 2006.

Karthik Muralidharan and Nishith Prakash. Cycling to school: increasing secondary school enrollment for girls in India. *American Economic Journal: Applied Economics*, 9 (3):321–50, 2017.

Karthik Muralidharan, Abhijeet Singh, and Alejandro J Ganimian. Disrupting education? experimental evidence on technology-aided instruction in india. *American Economic Review*, 109(4):1426–60, 2019.

Christophe Jalil Nordman, Smriti Sharma, and Naveen Sunder. Here comes the rain again: Productivity shocks, educational investments and child work. *Economic Development and Cultural Change, Forthcoming*. doi: 10.1086/713937.

Olayinka A Onasoga, Joel A Afolayan, and Bukola D Oladimeij. Factors influencing utilization of antenatal care services among pregnant women in ife central lga, osun state, nigeria. *Advances in Applied Science Research*, 3(3):1309–1315, 2012.

Philip Oreopoulos, Marianne E Page, and Ann Huff Stevens. The intergenerational effects of compulsory schooling. *Journal of Labor Economics*, 24(4):729–760, 2006.

Una Okonkwo Osili and Bridget Terry Long. Does female schooling reduce fertility? Evidence from nigeria. *Journal of Development Economics*, 87(1):57–75, 2008.

Emily Oster and Rebecca Thornton. Menstruation, sanitary products, and school attendance: Evidence from a randomized evaluation. *American Economic Journal: Applied Economics*, 3(1):91–100, 2011.

Dillee Prasad Paudel, Baburao Nilgar, and Manisha Bhandankar. Determinants of postnatal maternity care service utilization in rural belgaum of Karnataka, India: A community based cross-sectional study. *International Journal of Medicine and Public Health*, 4 (1), 2014.

Nicola Persico, Andrew Postlewaite, and Dan Silverman. The effect of adolescent experience on labor market outcomes: The case of height. *Journal of Political Economy*, 112(5): 1019–1053, 2004.

Marc Piopiunik. Intergenerational transmission of education and mediating channels: Evidence from a compulsory schooling reform in Germany. *The Scandinavian Journal of Economics*, 116(3):878–907, 2014.

Nitya Rao. Total literacy campaigns: a field report. *Economic and Political Weekly*, pages 914–918, 1993.

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.

Bruce Sacerdote. The nature and nurture of economic outcomes. *American Economic Review*, 92(2):344–348, 2002.

Bruce Sacerdote. How large are the effects from changes in family environment? a study of Korean American adoptees. *The Quarterly Journal of Economics*, 122(1):119–157, 2007.

Shanika Samarakoon and Rasyad A Parinduri. Does education empower women? Evidence from Indonesia. *World Development*, 66:428–442, 2015.

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

Mary A Silles. The intergenerational transmission of education: New evidence from adoptions in the usa. *Economica*, 84(336):748–778, 2017.

Bibha Simkhada, Edwin R van Teijlingen, Maureen Porter, and Padam Simkhada. Factors affecting the utilization of antenatal care in developing countries: systematic review of the literature. *Journal of Advanced Nursing*, 61(3):244–260, 2008.

Abhijeet Singh. Private school effects in urban and rural India: Panel estimates at primary and secondary school ages. *Journal of Development Economics*, 113:16–32, 2015.

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. *Demography*, 56(6):2123–2146, 2019.

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

Sebastian Vollmer, Christian Bommer, Aditi Krishna, Kenneth Harttgen, and SV Subramanian. The association of parental education with childhood undernutrition in low- and middle-income countries: comparing the role of paternal and maternal education. *International journal of Epidemiology*, 46(1):312–323, 2016.

Huan Wang, James Chu, Prashant Loyalka, Xin Tao, Yaojiang Shi, Qinghe Qu, Chu Yang, and Scott Rozelle. Can school counseling reduce school dropout in developing countries? *Rural Education Action Program Working Paper No*, 275, 2014.

Kin Bing Wu, Venita Kaul, and Deepa Sankar. Education and development-the quiet revolution-how India is achieving universal elementary education. *Finance and Development-English Edition*, 42(2):29–31, 2005.

Ji Yan. Maternal pre-pregnancy bmi, gestational weight gain, and infant birth weight: A within-family analysis in the united states. *Economics & Human Biology*, 18:1–12, 2015.

Joanne Yoong. The impact of economic resource transfers to women versus men: a systematic review. 2012.

Kathryn M Yount, AliceAnn Crandall, and Yuk Fai Cheong. Women's age at first marriage and long-term economic empowerment in Egypt. *World Development*, 102:124–134, 2018.

Table 1: **Effect of DPEP on education of direct beneficiaries**

DLHS - 3	**Male**				**Female**			
	Edu years	Primary school (=1)	Ever School	Literate	Edu years	Primary school (=1)	Ever School	Literate
DPEP	0.84***	0.12***	0.10***	0.09***	0.90***	0.05***	0.09**	0.08
	[0.15]	[0.03]	[0.021]	[0.022]	[.17]	[.02]	[.04]	[.016]
Intent-to-treat effect	0.36	0.05	0.04	0.04	0.39	0.02	0.04	0.03
Total Obs	373381	373381	501641	501777	450032	450032	500714	500825
Bandwidth	6	7.5	6.1	6	6.9	8	7.7	8.2
Obs. in bandwidth	92332	108520	112484	108496	117237	130580	141555	144637

DLHS - 4	**Male**			**Female**		
	Edu years	Primary school (=1)	Ever School	Edu years	Primary school (=1)	Ever School
DPEP	0.75***	0.05***	0.08**	0.78***	0.08**	0.03
	[0.21]	[0.03]	[.04]	[0.23]	[.04]	[0.03]
Intent-to-treat effect	0.33	0.02	0.03	0.34	0.03	0.01
Total Obs	159234	159234	159234	164367	164367	164367
Bandwidth	5.8	3.1	3.2	3.3	3.3	3.4
Obs. in bandwidth	50431	34898	34898	36650	36650	37562

Notes: This table reports the IV estimate of the effect of DPEP, where the first stage is based on the discontinuous change in probability of DPEP around the average national female literacy. The first stage F-statistic is 142.6, and the sample for this table consists of individuals who were born in/between 1980 and 1995. The RD impact point estimates are constructed using the triangular kernel, local polynomial of order 2 and with the CER-optimal bandwidth selector bandwidth selector. Standard errors are robust-bias corrected and clustered at the district level. The table is based on data from rounds 3 and 4 of the District level Household and Facility Survey DLHS which were held in 2007-08 and 2012-13 respectively.

Table 2: **Intergenerational impact of DPEP on children of mother beneficiaries**

	Reading Score	Math Score	English Score
DPEP	.41***	.27***	.32**
	[.12]	[.09]	[.16]
First Stage coefficient	0.42***	0.41***	0.42***
First Stage F-statistic	108.7	103	122.4
Bandwidth	6.9	6.4	7
Obs. in bandwidth	93208	87064	53089
Control Mean	2.64	2.32	2.25
Total Obs	331314	330313	198635

Notes: The sample for this part of the analysis consists of children who satisfy the following conditions – (1) have mothers who were born between 1980 and 1995, and were thereby direct DPEP beneficiaries, (2) have fathers who were not direct DPEP beneficiaries, and (3) would themselves have started school after 2005 (after DPEP had been phased-out). The point estimates presented here are the coefficients from a fuzzy RD specification (akin to second-stage estimates from a 2SLS specification). Here the first stage is based on the discontinuous change in probability of DPEP around the average national female literacy. The first stage coefficient estimate is 0.43, and the estimates in the table are constructed using the triangular kernel, local polynomial of order 2 and with CER-optimal bandwidth selector. The specifications control for categorical variables for gender of the child, state and year of data collection. The standard errors are clustered at the district level. The bandwidth is expressed in terms of the running variable - district female literacy rate in 1991. The mean of the outcome among non-beneficiaries is reported (for observations within the bandwidth). Data Source: Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018.

Table 3: **Intergenerational impact of DPEP on children of father beneficiaries**

	Reading Score	Math Score	English Score
DPEP	.027	.07	.06
	[.09]	[.12]	[.09]
First Stage coefficient	0.43***	0.43***	0.42***
First Stage F-statistic	104.7	127.9	96.1
Bandwidth	6	7.3	6.6
Obs. in bandwidth	22026	25957	5842
Control Mean	2.36	2.21	2.26
Total Obs	118376	117884	28611

Notes: The sample for this part of the analysis consists of children who satisfy the following conditions – (1) have fathers who were born between 1980 and 1995, and were thereby direct DPEP beneficiaries, (2) have mothers who were not direct DPEP beneficiaries, and (3) would themselves have started school after 2005 (after DPEP had been phased-out). The point estimates presented here are the coefficients from a fuzzy RD specification (akin to second-stage estimates from a 2SLS specification). Here the first stage is based on the discontinuous change in probability of DPEP around the average national female literacy. The first stage coefficient estimate is 0.44, and the estimates in the table are constructed using the triangular kernel, local polynomial of order 2 and with CER-optimal bandwidth selector. The specifications control for categorical variables for gender of the child, state and year of data collection. The standard errors are clustered at the district level. The bandwidth is expressed in terms of the running variable - district female literacy rate in 1991. The mean of the outcome among non-beneficiaries is reported (for observations within the bandwidth). Data Source: Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018.

Table 4: **Mechanisms - Woman (Parent) & Child level**

	(1) Decision-Child	(2) Decision-Cook	(3) Decision-Purchases	(4) Beat-Bad Cook	(5) Beat-Neglect House
DPEP	0.08***	0.14***	0.08***	-0.09***	-0.05***
	[0.02]	[0.053]	[0.03]	[0.02]	[0.01]
Bandwidth	8.4	9	10.5	7.4	7.4
Obs. in bandwidth	4449	4826	5998	3960	3960
Control Mean	.69	.87	.61	.31	.34
First Stage F-stat	104.2	109.3	110.5	99.7	89.2
Total Obs	14297	14297	14297	14297	14297

	(1) School Fees	(2) Uniform/Books	(3) Tuition Fees	(4) Private School	(5) Homework Hours
DPEP	58.8	706.8**	156.3***	0.05	2.06*
	[46.3]	[328.6]	[52.8]	[0.05]	[1.16]
Bandwidth	9.1	8.6	12.2	10.6	10.6
Obs. in bandwidth	6727	6088	9415	8315	8009
Control Mean	1847.14	1728.39	528.99	.07	7
First Stage F-stat	108.2	106.1	110.5	99.7	89.2
Total Obs	19146	19323	18541	19571	18964

Notes: The sample consists of women who were born between 1980 and 1995, and children born to these women who would have started school in or after 2005. The RD point estimates are constructed using the triangular kernel, local polynomial of order 2 and with CER-optimal bandwidth selector. The standard errors are robust-bias corrected and are clustered at the district level. Decision-Cook and Decision-Purchases are dummy variables that take a value of one if the woman reported being involved in decisions related to cooking and purchases made in the household. *Beat-Bad Cook* and *Beat-Neglect House* are categorical variables that take a value of one if woman reported that it was common for women in their community to be beaten up in cases when she was a bad cook or neglected household work respectively. School Fees, Uniform/Books and Tuition fees refer to variables that measure the expenditure on these categories made on the woman's children. Private School is a dummy that takes a value of one if the child goes to private school, and homework hours refers to the amount of time the child spent doing homework. Data Source: Indian Human Development Survey (IHDS) 2005 and 2012 rounds.

Table 5: **Mechanisms - Woman (Parent) Level**

VARIABLES	(1) Marriage Age	(2) Age at First Birth	(3) Contraceptive Use	(4) Any ANC	(5) # ANC visits	(6) Delivery - Formal	(7) Any PNC
DPEP	0.64***	0.33**	.06***	0.07**	0.22**	0.06***	.06***
	[0.25]	[0.14]	[.02]	[.03]	[0.12]	[.02]	[.023]
Bandwidth	5.9	9.7	7.5	10.2	10.2	11	11.3
Obs. in bandwidth	34950	45270	26642	36864	29557	41798	44175
Control Mean	18.2	19.62	.35	.80	3.99	.52	.51
Total Obs	166070	136859	108817	109973	82864	109976	109975

Notes: The sample consists of women who were born between 1980 and 1995. The RD point estimates are constructed using the triangular kernel, local polynomial of order 2 and with CER-optimal bandwidth selector. The standard errors are robust-bias corrected and are clustered at the district level. Marriage age refers to the age at marriage (in years), Age at first birth refers to age when the woman had her first child (in years), Contraceptive use is a dummy that takes a value of one if the women reported using contraceptives, *Any ANC* is a categorical variable that takes a value of one if the woman accessed any ANC facilities during the last pregnancy, # ANC visits refers to the number of ANC visits made during the last pregnancy, *IFA Taken* is a categorical variable that takes a value of one if the woman reported taking Iron and Folic Acid (IFA) tablets during the last pregnancy, *Delivery-Formal* is a categorical that takes a value of one if the woman reported giving birth in a formal health facility and *Any PNC* refers to a dummy that takes a value of one if the woman reported using any Post Natal Care (PNC) facilities. Data Source: District Level Household and Facility Survey (DLHS) data Round 3 (2007-08).

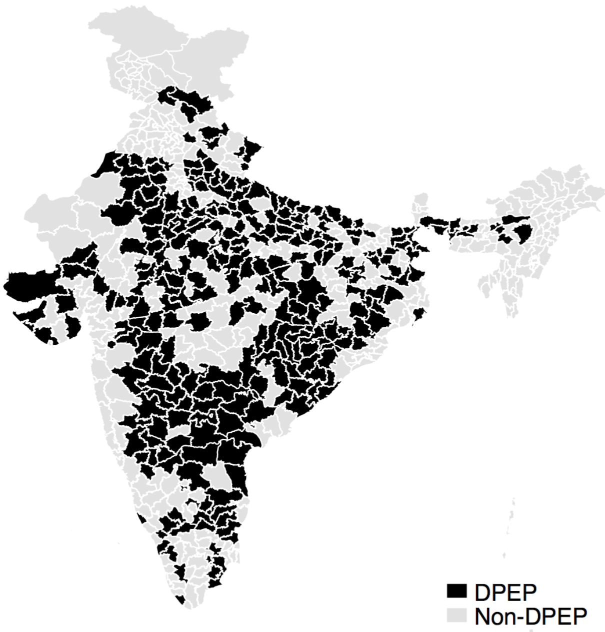

Figure 1: **Map of DPEP and non-DPEP districs.**

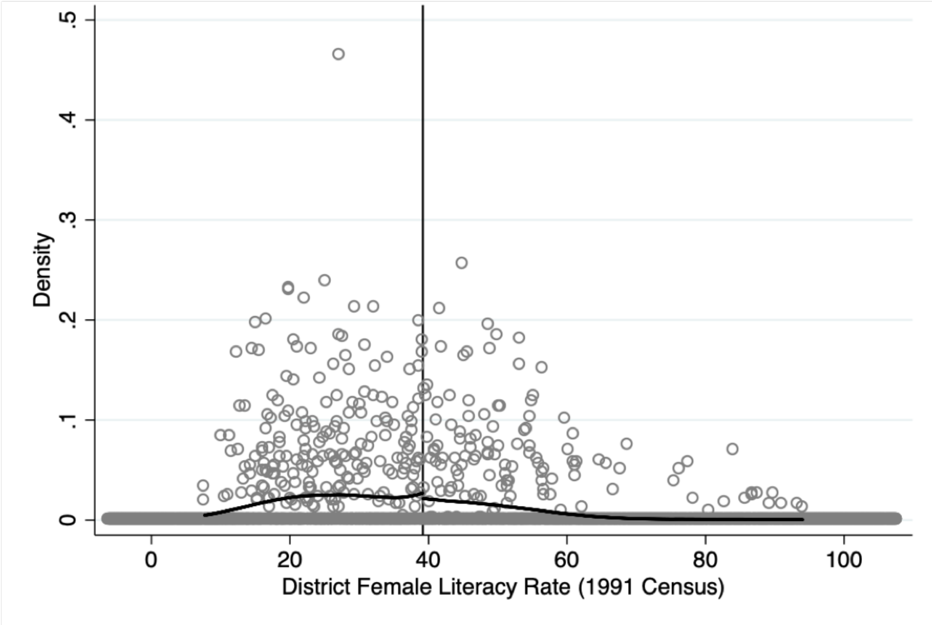

Figure 2: **McCrary Density test**

Notes: Conducted based on procedure outlined in McCrary [2008]. The associated test statistic is 0.003, and the p-value is 0.72. Based on author's calculations using the Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018.

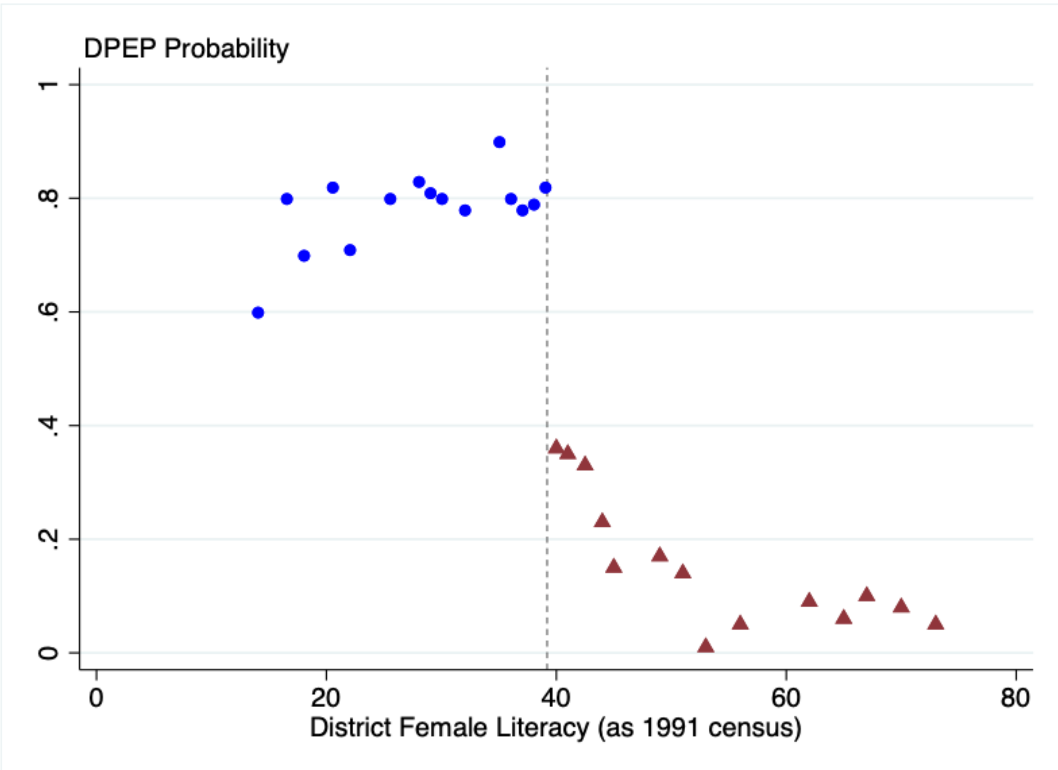

Figure 3: **Probability of Receiving DPEP Programme.**

Notes: The graph illustrates the discontinuity of treatment assignment at the cutoff of 39.2 percent in terms of District Female Literacy Rate. Data on district female literacy rate is from the 1991 Census.

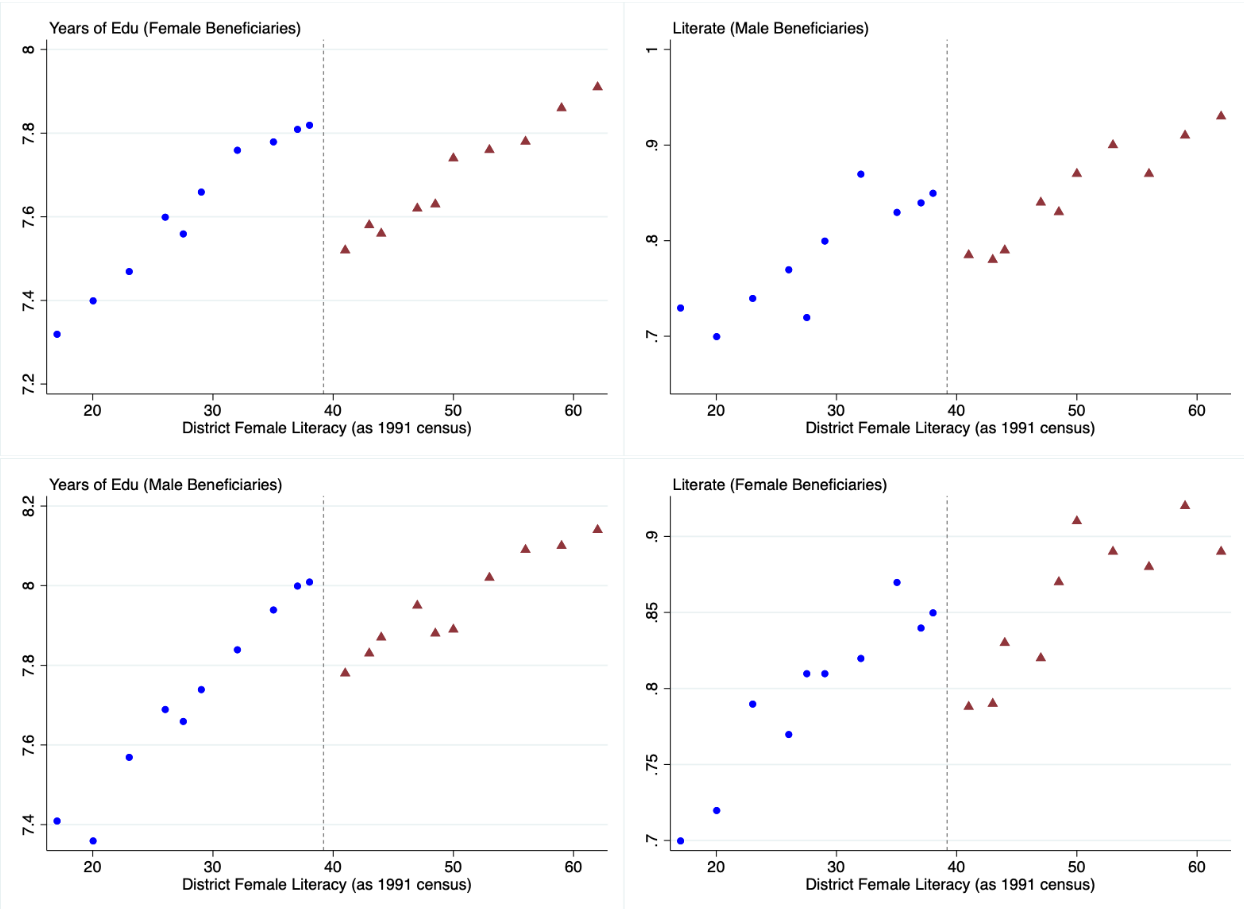

Figure 4: **Impact of DPEP on direct beneficiaries**

Notes: The graph presents the differences in educational outcomes for male and female direct DPEP beneficiaries. The sample consists of individuals who were born between 1980 and 1995. Data Source: District Level Household and Facilities Survey (rounds 3 & 4).

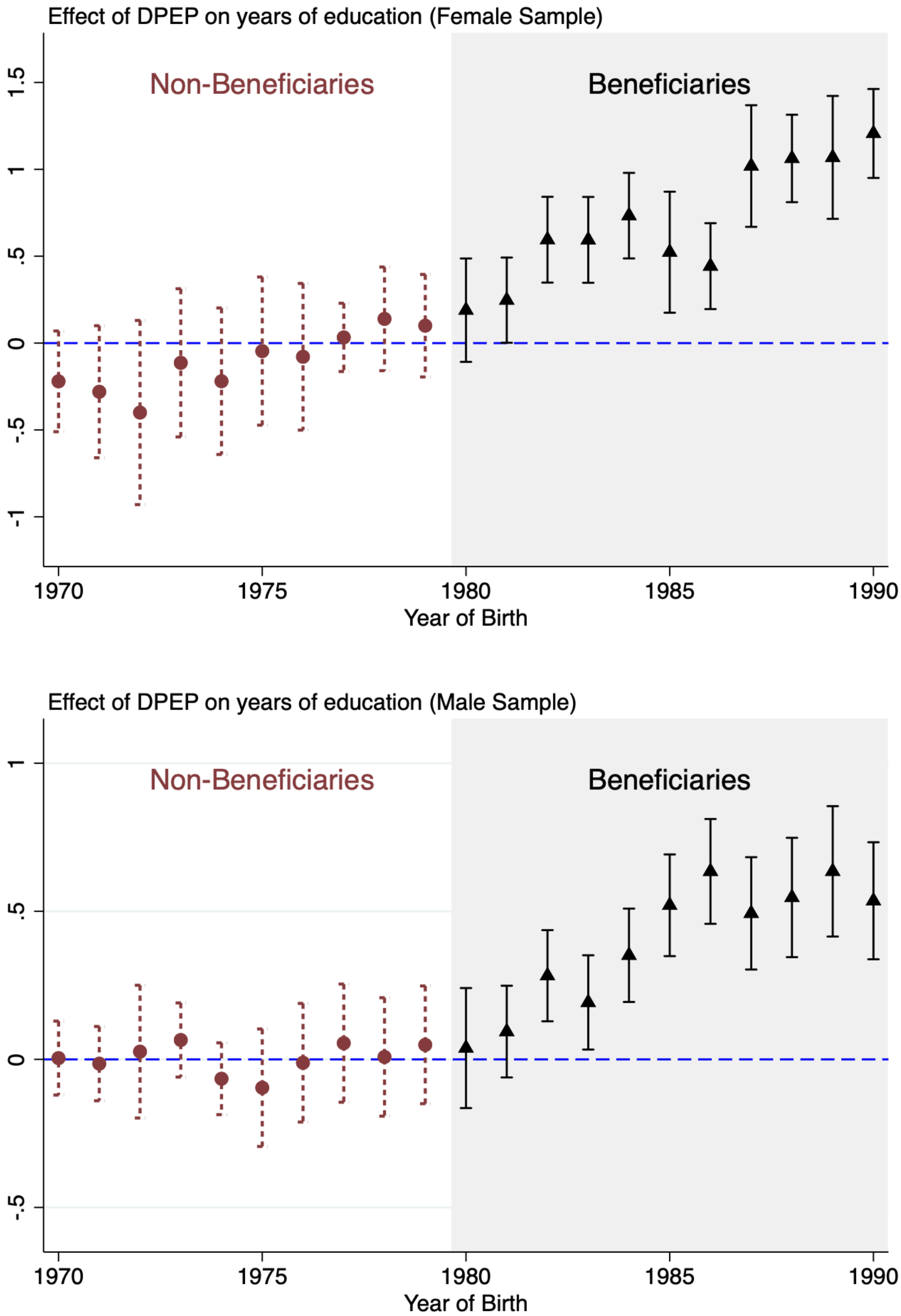

Figure 5: **Difference between DPEP and non-DPEP regions- cohortwise**

Notes: The plot contains the coefficients of DPEP dummy (=1 for districts that received the policy) interacted with year of birth in a specification where the outcome variable is years of education, and the controls include survey year fixed effects, state fixed effects and state-specific time trends. The sample for the analysis is individuals born between 1970 and 1990. The analysis is done separately for female and male sub-samples. The dashed vertical line represents the distinction between the DPEP beneficiary group (born after 1980) and the non-beneficiary group (born before 1980). The 95 percent confidence intervals are presented in the figure.

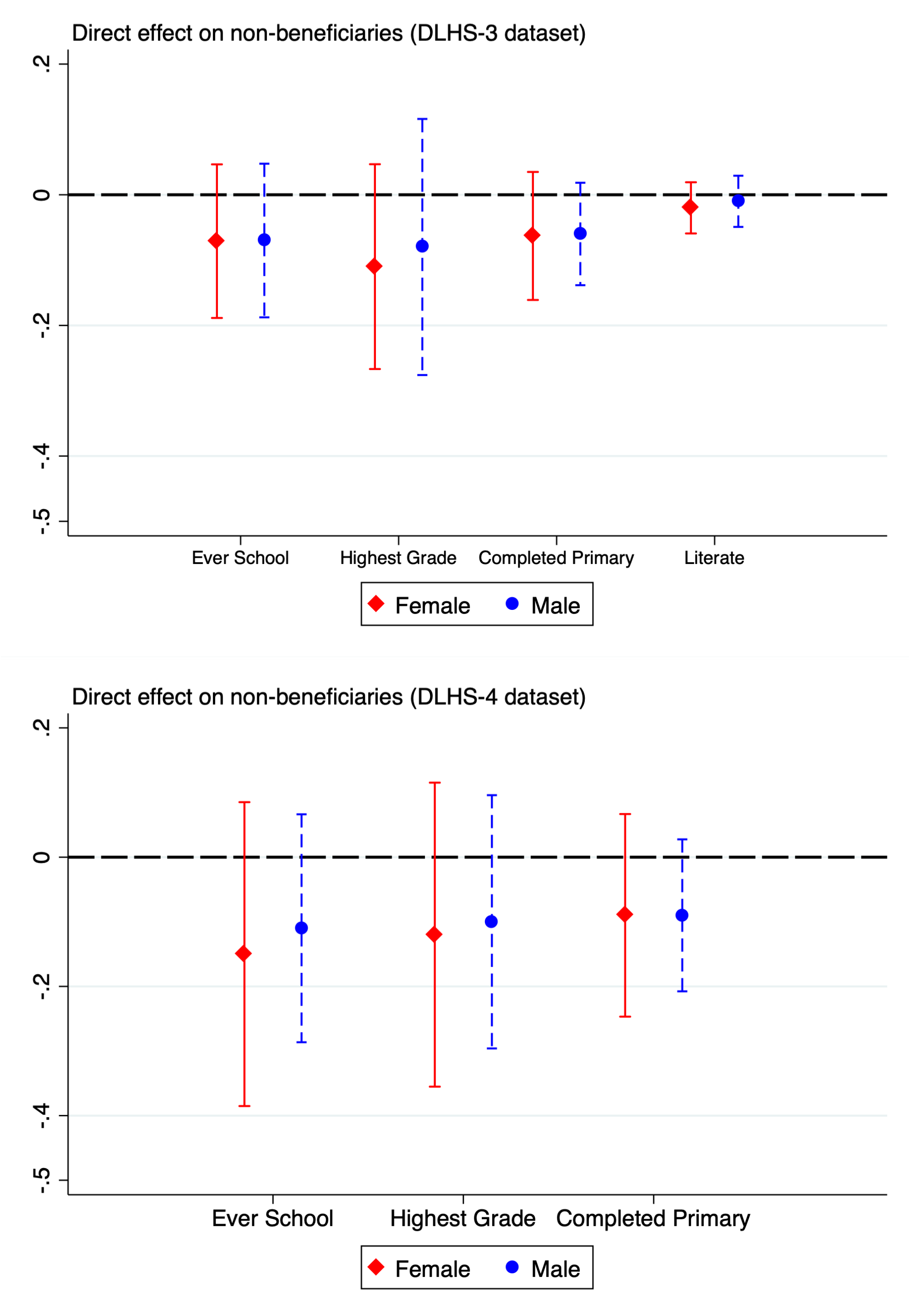

Figure 6: **Falsification Check – DPEP impact on non-beneficiaries**

Notes: The sample consists of individuals born between 1968 and 1975. The analysis is conducted separately for females and males, and the point estimates are constructed using the CER-optimal bandwidth selector. The standard errors are clustered at the district level, and the 95 percent confidence intervals are presented. Data: DLHS rounds 3 and 4.

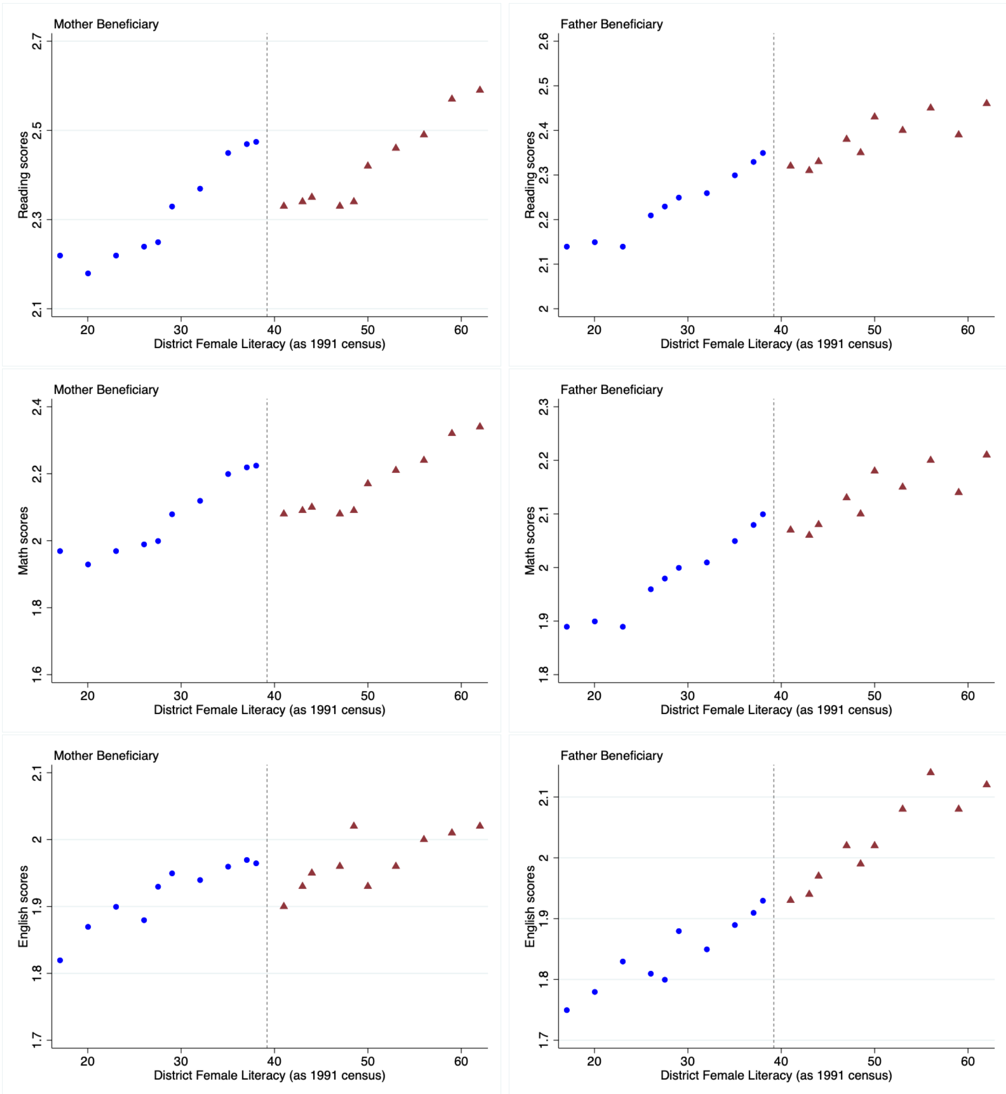

Figure 7: **Impact of DPEP on intergenerational test scores.**

Notes: The graphs in the top, middle and bottom panel correspond to reading, math and english scores respectively. Based on author's calculations using the Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018.

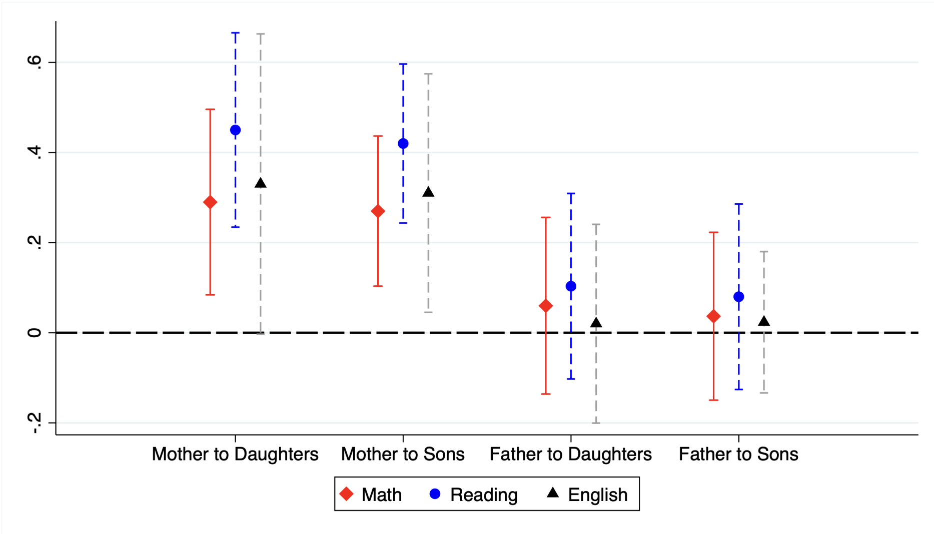

Figure 8: **Intergenerational DPEP impact – Genderwise results**

Notes: The analysis sample for the "Mothers to Daughters" and the "Mothers to Sons" consists of children who satisfied two conditions - (1) their mothers were born between 1980 and 1995, and (2) they started school after 2005 (end of DPEP). Analogously, The analysis sample for the "Fathers to Daughters" and the "Fathers to Sons" consists of children who satisfied two conditions - (1) their fathers were born between 1980 and 1995, and (2) they started school after 2005 (end of DPEP). The point estimates are constructed using the triangular kernel, local polynomial of order 2 and with CER-optimal bandwidth selector. All specifications control for categorical variables for state and year of data collection. The standard errors are clustered at the district level, and the 95 percent confidence intervals are presented. All the score variables run from zero to four. Data Source: Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018.

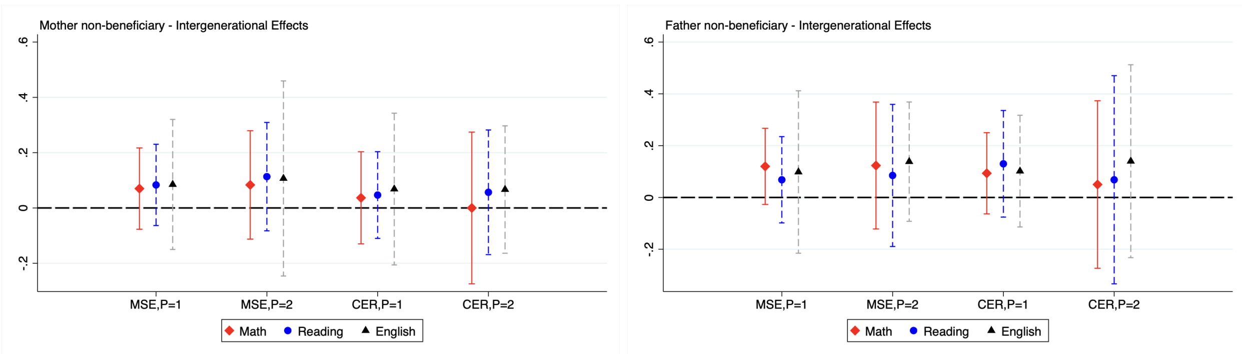

Figure 9: **Falsification Check – Intergenerational DPEP impact on non-beneficiaries**

Notes: The sample consists of children who satisfied two conditions – (1) whose mothers (fathers) were 18-25 years of age in 1993 (born between 1968 and 1975) and (2) they started school after 2005 (end of DPEP). All specifications control for categorical variables for gender of the child, state and year of data collection. The standard errors are clustered at the district level, and the 95 percent confidence intervals are presented. Data Source: Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018.

APPENDIX A – Changes in District Boundaries

India has witnessed substantial administrative decentralization over the past few decades, during which the number of districts in the country has increased from 466 (in 1991) to 640 (in 2011). Given that the DPEP programme was implemented at the district level, it is crucial that I be able to link current district definitions to districts boundaries that existed during the program years (1993-2004)[34]. This is important because the parent cohort went to school in old districts (1993-2000), and hence, whether or not they were received the benefits of DPEP would depend on their district of residence during program implementation.

Based on the discussion of historical changes in district boundaries in Kumar and Somanathan, 2009, I map districts in 2001 to their parent districts in 1991. While Kumar and Somanathan, 2009 only cover district changes that took place until 2001, I extend their analysis to similarly match districts that were created in subsequent years (until 2011). In doing so, I follow these steps. In some cases, multiple districts were created from a single parent district, and so I assign the treatment status of the parent district to all the new districts. In other cases, several districts were combined to form a new big district. Here, if all the parent districts had the same treatment status, I assign the same status to the new district. However, there are cases in which the DPEP treatment status of the parent districts differ. If so, if more than 50 percent (or more) of the population of the new district comes from parent districts of a certain treatment status, I assign this treatment status to the new district. I use analogous rules in assigning district characteristics (such as the 1991 district female literacy rates) to the newly created districts.

[34] India is administratively split up first into states, which are then split up into districts. These are akin to counties in the US.

APPENDIX B – School Infrastructure in the years 1993 to 2004

I test for differences in the school infrastructure at three points in time: in 1993 (Pre-DPEP), between 1993 & 2004 (DPEP years) and in 2005 (end of DPEP). This would help understand the changes in school quantity and quality that happened before, during and after DPEP implementation.

School Quality

I use school-level census data from DISE to examine several school quality measures. These include physical infrastructure (classrooms, toilets, electricity), teacher qualification, school oversight (inspection visits) and grants/incentives received by the school (funding received/spent). Put together, these arguably provide a comprehensive overview of the amenities/resources that a school possesses and is a wide enough array of indicators so as to encompass enough aspects of school quality. The results in Table A2 show that there were no differences in school quality across the RD cutoff at the start of the DPEP policy (in 1993)[35] and those constructed during the DPEP years (between 1993 & 2004)[36].

School Quantity

I conduct a similar analysis for the quantity of school infrastructure. In table A1 (panel B), I show that, as expected, there was enhanced school construction in DPEP regions during the policy years. Comparing this to results in panel A and C of table A1, it suggests that the enhanced school construction happened only during DPEP years, and not at the beginning of DPEP (panel A, table A1) and after DPEP had been phased out (panel C, table A1). Put together, this suggests that there were more schools constructed in DPEP

[35] Ideally, to do this one would have used data on these measures from the year 1993-94. Since such detailed data is unavailable for that time period, I use data from 2005 for schools that were constructed before 1993. This estimation would be valid if there were no systematic differences in upgrades/improvements in schools constructed before 1993 in districts around the programme cutoff. There is no reason to believe that this would be the case.

[36] Ideally, I would want to compare the schools built under the DPEP programme to other schools constructed in this time period to show that the DPEP schools were no different from the other schools. But the dataset does not identify the schools specifically built under this programme. So I compare all schools constructed in this period in districts around the cutoff. Given that DPEP was the flagship government programme for that period, it can be argued that any differences in school quality should be captured in this setup.

districts only when the programme was ongoing, and not in any other time period. Additionally, results in panel A of table A1 suggest that there were fewer schools in DPEP regions before the policy was implemented – this is consistent with the fact that DPEP was implemented in lower female literacy regions (which plausibly is correlated with diminished school availability). Therefore this implies that there were fewer schools to start with in DPEP regions, but then these regions experienced enhanced rates of school construction during DPEP (and at no other time). This leads to another question - were there any differences in the number of schools between DPEP and non-DPEP districts at the end of DPEP? To investigate this, I examine the stock of schools (total, government and private) in 2006 (when the DPEP policy had ended) – I find that DPEP districts had more schools at the end of DPEP but the difference is not statistically significant.

The pattern described in the regression table can also be observed using figures. The upper panel of figure A16 shows that more schools were constructed (per-capita) in DPEP regions. The bottom panel of this figure presents the yearwise difference in number of government schools constructed during the DPEP period - it indicates that there was a higher number of government schools constructed at the height of the DPEP policy (between 1993 and 2000), and this declined as the policy was phased out in the latter years. In addition to this, I also show that there were no differences in the rate of school construction (figure A2) and the total number of schools (figure A3) in the post-DPEP period (between 2006 to 2012) – this is the period when the second-generation beneficiaries were attending school.

Put together, the lack of differences in school quality and quantity in the post-DPEP period alleviates concerns about the positive intergenerational learning effects being driven by school infrastructure differences experienced by the children of direct beneficiaries.

APPENDIX C – DPEP effect on Direct Beneficiaries using Difference-in-Differences

In a supplementary analysis I estimate DPEP effects on direct beneficiaries using a difference-in-differences approach. Here, I leverage two sources of variations – whether a district received DPEP or not, and exposure to the policy depending on the age of the individual during policy years. As discussed earlier, DPEP was targeted towards primary school-going age children, and hence the beneficiary group was between 5 and 14 years of age during the policy years (1994 to 2000). This implies that the oldest beneficiary cohort (14 year olds in 1994) would have been born in 1980, and people born before this would have an extremely low likelihood of being affected by DPEP. I use this intuition to estimate the following model:

$$Y_{isdt} = \alpha + \sum_{j=1970}^{1990} \delta_j \text{ cohort }_j * DPEP_{sd} + \sum_s \theta_s \text{ state}_s * trend + \mu X_{isdt} + \gamma_s + \gamma_t + \epsilon_{isdt} \tag{5}$$

Here Y_{isdt} refers to the outcome of interest for person i in district d of state s who was surveyed in year t, cohort_j is a categorical variable that takes a value of one for individuals born in year j, $DPEP_{sd}$ is a categorical variable for whether district d in state s received the DPEP policy or not, X_{isdt} refers to control covariates (categorical variables for for caste and religion), state fixed effects (γ_s), year of survey fixed effects (γ_t), and state-specific time trends ($\text{state}_s * trend$). Standard errors are clustered at the district-level. This specification is akin to the analysis in Azam and Saing, 2017.

In this setup, the coefficients of interest are represented by δ_j, which consists of one coefficient for each cohort (j) born between 1970 and 1990[37]. I run this specification separately for female and male sub-samples, and the results are presented in figure 5. Here, I plot the coefficients on the interactions between DPEP status and birth-year dummies for each year between 1970 and 1990. The graphs suggest the following: (1) the positive effect of DPEP is present for cohorts born in or after 1980, and is largely concentrated in

[37] This provides an estimate of the treatment effect on the treated (TOT).

the years after 1982[38], (2) there is no significant effect of DPEP on individuals born before 1980, and (3) the aforementioned patterns are observed for both females (top panel of figure 5) and males (bottom panel of figure 5). These findings are in line with our expectations that the beneficiaries (1980 to 1990 birth cohorts) show positive programme effects on years of education, while the non-beneficiaries (1970 to 1979 birth cohorts) experience null effects.

To be able to conduct this kind of a diff-in-diff analysis, it is important that there are parallel trends in the outcomes in the pre-reform period. This is partially shown by the fact that the coefficients in the pre-reform period (1979 and before) in are statistically insignificant points to parallel trends prior to DPEP initiation. Further, I show that the mean years of education for both males and females born in or before 1979 (non-beneficiary group) is similar across both DPEP and non-DPEP districts (figure A17).

I further add to this analysis by using a similar diff-in-diff setup to explore intergenerational effects.

$$Y_{isdt} = \alpha + \delta_j \text{ Mother of eligible age }_i * DPEP_{sd} + \mu X_{isdt} + \gamma_s + \gamma_t + \epsilon_{isdt} \tag{6}$$

In this model I leverage two sources of variation: (i) *DPEP District* (whether a district received the DPEP policy or not), and (ii) *Mother of eligible age* - whether the mother of the child was of school-going age or not during the DPEP years (beneficiary group was between 5 and 14 years of age between the years 1994 and 2000). When examining the effects of having a *Father of eligible age*, I define the corresponding variable similarly– whether the father of the child was between 5-14 years of age between 1994 and 2000. In all specifications I control for gender of the child, child age fixed effects, state fixed effects and year of data collection, and the standard errors are clustered at the district level (as the treatment was allocated at this administrative level).

[38]One would expect that individuals 12-14 years in 1994 (1980-82 birth cohorts) would have benefitted the least due to the policy as they were close to finishing primary school. A higher beneficial effect would be expected among those who had more years of primary school left.

I find that the main findings (tables 2 and 3) are consistent with what we observe using this specification in table A6 – mothers' exposure to DPEP leads to improved learning outcomes for children whereas father's exposure leads to no statistically significant effect on learning. The samples used here are the same as those in tables 2 and 3 – this means that I can compare the coefficients we get here with those in tables 2 and 3. I find that they are similar in magnitude and there are no statistically significant differences (as the confidence intervals coincide).

APPENDIX TABLES & FIGURES

Table A1: **School Construction in and around DPEP-years**

	Panel A: Schools constructed before 1993 (pre-DPEP)					
	All of Schools			Schools per 1000 Population		
	All Schools	Government	Private	All Schools	Government	Private
DPEP	-143.7	-104.2	-50.77	-0.05	-0.04	-0.01
	[343.9]	[301.6]	[83.2]	[0.28]	[0.49]	[0.03]

	Panel B: Schools constructed between 1993 and 2005 (during DPEP)					
	All Schools			Schools per 1000 Population		
	All Schools	Government	Private	All Schools	Government	Private
DPEP	413.8**	258.1***	50.77	0.31**	0.21***	0.08
	[173.3]	[98.6]	[83.2]	[0.14]	[0.09]	[0.07]

	Panel C: Schools constructed after 2005 (post-DPEP years)					
	All Schools			Schools per 1000 Population		
	All Schools	Government	Private	All Schools	Government	Private
DPEP	27.1	15.9	10.77	0.03	0.04	0.01
	[33.9]	[35.6]	[13.2]	[0.05]	[0.09]	[0.03]
Total Obs	442	442	442	440	440	440

	Panel D: Total number of schools in 2006 (after end of DPEP)					
	All Schools			Schools per 1000 Population		
	All Schools	Government	Private	All Schools	Government	Private
DPEP	21.9	5.4	16.7	0.01	0.04	0.11
	[23.6]	[41.3]	[19.2]	[0.06]	[0.09]	[0.10]
Total Obs	442	442	442	440	440	440

Notes: Based on author's calculations using the District Information on System of Education (DISE) data. The point estimates are constructed using the CER-optimal bandwidth selector.

Table A2: **Impact of DPEP on School Quality**

	(1) Total Classrooms	(2) Girls Toilets	(3) Boys Toilets	(4) Electricity	(5) # Male Teachers	(6) # Female Teachers	(7) # Graduate Teachers
	Panel A - Schools constructed before 1993						
DPEP	0.11	0.16	0.06	-0.06	0.07	0.19	0.51
	[.13]	[.37]	[.35]	[.13]	[0.35]	[0.18]	[1.1]
Bandwidth	1.5	.7	.7	.4	.4	.5	.6
Total Obs	654545	654545	654545	654545	654545	654545	654545
	Panel B - Schools constructed between 1993 and 2005						
DPEP	0.16	-0.04	0.02	0.032	-0.38	0.06	0.044
	[0.11]	[0.06]	[0.13]	[0.18]	[.63]	[0.1]	[0.05]
Bandwidth	1.5	2	5.2	1.3	.9	1	1.4
Total Obs	347703	347703	347703	347703	347703	347703	347703
	Panel C - Schools constructed after 2005						
DPEP	0.01	0.03	0.04	0.28	0.17	0.07	0.05
	[0.15]	[.06]	[0.05]	[0.2]	[0.27]	[0.07]	[0.09]
Bandwidth	1.6	2.4	2	1.5	1.9	2.1	1.1
Total Obs	188685	188685	188685	188685	188685	188685	188685

	(1) # Visits BRC	(2) # Visits CRC	(3) Any SMC	(4) Devt. grant received	(5) Devt. grant spent	(6) Maintenance grant received	(7) Maintenance grant spent
	Panel D - Schools constructed before 1993						
DPEP	0.09	0.13	0.16	-10.16	18.03	34.08	11.05
	[.13]	[.17]	[.25]	[20.3]	[20.05]	[46.08]	[19.1]
Bandwidth	1.5	.5	.3	.5	.5	.6	1
Total Obs	654545	654545	654545	654545	654545	654545	654545
	Panel E - Schools constructed between 1993 and 2005						
DPEP	0.12	-0.04	0.01	22.32	17.18	48.06	8.44
	[0.1]	[0.05]	[0.08]	[40.18]	[15.63]	[70.1]	[17.05]
Bandwidth	1.5	1.3	.6	1.5	1.6	2.2	2.1
Total Obs	347703	347703	347703	347703	347703	347703	347703
	Panel F - Schools constructed after 2005						
DPEP	0.15	-0.13	-0.14	33.78	20.17	60.12	11.05
	[0.14]	[.16]	[0.12]	[51.6]	[38.7]	[78.09]	[9.07]
Bandwidth	1.6	3.6	1.9	5.2	4.6	6	3.4
Total Obs	188685	188685	188685	188685	188685	188685	188685

Notes: Based on author's calculations using the District Information on System of Education (DISE) district level data for the year 2014. The RD point estimates are constructed using a local polynomial of order 2 and with CER-optimal bandwidth selector. The standard errors are robust-bias corrected and are clustered at the district level.

Table A3: **Intergenerational impact of DPEP on children when both parents are beneficiaries**

	Reading Score	Math Score	English Score
DPEP	.27**	.24***	.14**
	[.09]	[.07]	[.07]
Bandwidth	5.2	5.4	5.9
Obs. in bandwidth	167535	172075	82899
Control Mean	2.12	1.94	1.88
Total Obs	906736	903846	394469

Notes: The sample consists of children who satisfy two conditions - (1) their mother and father (both) were born between 1980 and 1995, and (2) they started school after 2005 (end of DPEP). The point estimates are constructed using the triangular kernel, local polynomial of order 2 and with one common CER-optimal bandwidth selector. All specifications control for categorical variables for gender of the child, state and year of data collection. The standard errors are clustered at the district level. The bandwidth is expressed in terms of the running variable - district female literacy rate in 1991. The mean of the outcome among non-beneficiaries is reported (for observations within the bandwidth). Data Source: Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018.

Table A4: **Summary Statistics**

	Full Sample			**Non-DPEP**			**DPEP**		
	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.
ASER Data – Mother Beneficiaries									
Reading score (0-4)	331314	2.6	1.5	187750	2.5	1.5	143564	2.6	1.5
Mathematics score (0-4)	330313	2.3	1.3	187209	2.2	1.3	143104	2.3	1.3
English score (0-4)	198635	2.1	1.5	112590	2.0	1.5	86045	2.2	1.5
Child female (=1)	332152	0.5	0.5	188224	0.5	0.5	143928	0.5	0.5
Child age (in years)	331665	9.9	3.1	187968	9.9	3.1	143697	9.9	3.1
ASER Data – Father Beneficiaries									
Reading score (0-4)	118376	2.2	1.5	69612	2.1	1.5	48764	2.4	1.4
Mathematics score (0-4)	117884	2.0	1.3	69340	1.9	1.3	48544	2.2	1.3
English score (0-4)	28611	1.9	1.5	16804	1.7	1.4	11807	2.1	1.5
Child female (=1)	118676	0.5	0.5	69798	0.5	0.5	48878	0.5	0.5
Child age (in years)	118380	8.6	2.2	69663	8.5	2.2	48717	8.6	2.2

Table A5: Potential mechanisms and their contribution

Outcome: Reading score

	(1) Contribution (%)	(2) DPEP on channel	(3) Channel on outcome	(4) DPEP on outcome
Marriage Age	2.53%	0.64	0.017	0.43
Age at First Birth	1.53%	0.33	0.02	0.43
Delivery-Formal	1.95%	0.06	0.14	0.43
Any PNC	3.07%	0.06	0.22	0.43
Decision - Child	5.02%	0.08	0.27	0.43
Decision - Cook	3.26%	0.14	0.1	0.43
Decision -Purchase	2.05%	0.08	0.11	0.43
Uniform/Books	12.03%	706.8	0.0000732	0.43
Tuition Fees	9.52%	156.3	0.000262	0.43

Outcome: Math score

	Contribution (%)	DPEP on channel	Channel on outcome	DPEP on outcome
Marriage Age	3.25%	0.64	0.0137	0.27
Age at First Birth	2.02%	0.33	0.0165	0.27
Delivery-Formal	3.05%	0.06	0.1372	0.27
Any PNC	3.09%	0.06	0.139	0.27
Decision - Child	8.59%	0.08	0.29	0.27
Decision - Cook	4.51%	0.14	0.087	0.27
Decision -Purchase	3.85%	0.08	0.13	0.27
Uniform/Books	16.54%	706.8	0.0000632	0.27
Tuition Fees	8.80%	156.3	0.000152	0.27

Outcome: English score

	Contribution (%)	DPEP on channel	Channel on outcome	DPEP on outcome
Marriage Age	2.48%	0.64	0.0124	0.27
Age at First Birth	1.17%	0.33	0.0113	0.27
Delivery-Formal	2.63%	0.06	0.14	0.27
Any PNC	3.00%	0.06	0.16	0.27
Decision - Child	3.00%	0.08	0.12	0.27
Decision - Cook	4.03%	0.14	0.092	0.27
Decision -Purchase	0.75%	0.08	0.03	0.27
Uniform/Books	12.72%	706.8	0.0000576	0.27
Tuition Fees	11.72%	156.3	0.00024	0.27

Notes: Column two presents the LATE estimates of the effect of DPEP on different mechanisms (from tables 5 and 4), column 3 presents the association between the mechanism and outcome of interest, while column 4 presents the LATE estimate of the effect of DPEP on the outcome of interest (based on table 2). Column 1 presents the relative contribution of each mechanism to the corresponding outcome, and is calculated by dividing column 2 by the product of columns 3 and 4.

Table A6: Effect of DPEP on learning outcomes using a diff-in-diff specification

	Reading	Math	English
Panel A			
DPEP District * Mother of eligible age	0.35**	0.25**	0.25*
	(0.17)	(0.13)	(0.14)
Total Observations	331,314	330,313	198,635
Panel B			
DPEP District * Father of eligible age	0.04	0.12	0.05
	(0.13)	(0.14)	(0.12)
Total Observations	118,376	117,884	28,611

Notes: Data Source: Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018. DPEP District refers to whether a district received the DPEP policy or not, and *Mother of eligible age* refers to whether the mother of the child was of school-going age (between 5 and 14 years of age) during policy years (between 1994 and 2000). The variable *Father of eligible age* is similarly defined. All specifications control for categorical variables for gender of the child, child age fixed effects, state fixed effects and year of data collection. The standard errors are clustered at the district level.

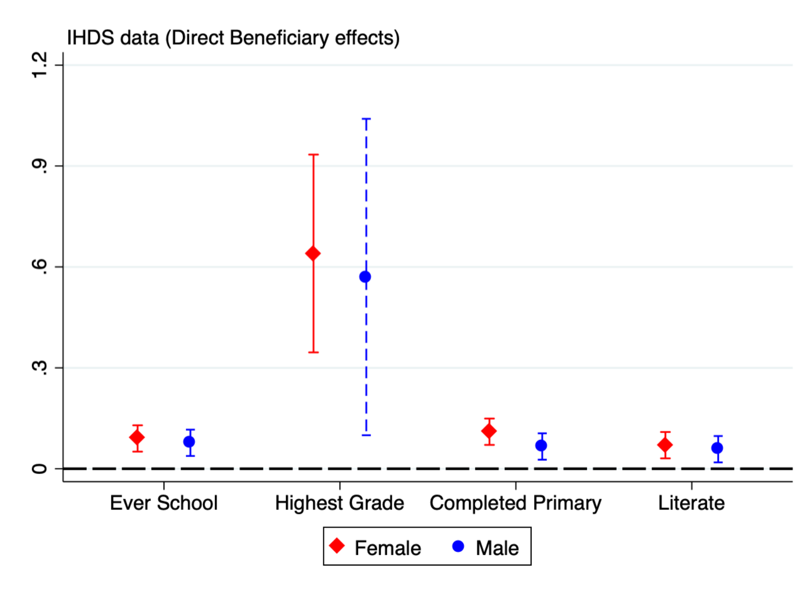

Figure A1: **DPEP impact on education (using IHDS data)**

Notes: The sample for this analysis consists of individuals who were born between 1980 and 1995. The impact point estimates are constructed using the triangular kernel, local polynomial of order 2 and with the CER-optimal bandwidth selector bandwidth selector. All the specifications control for categorical variables for caste, religion, and state fixed effect. Standard errors are robust and clustered at the district level. Data source: IHDS Rounds 1 and 2.

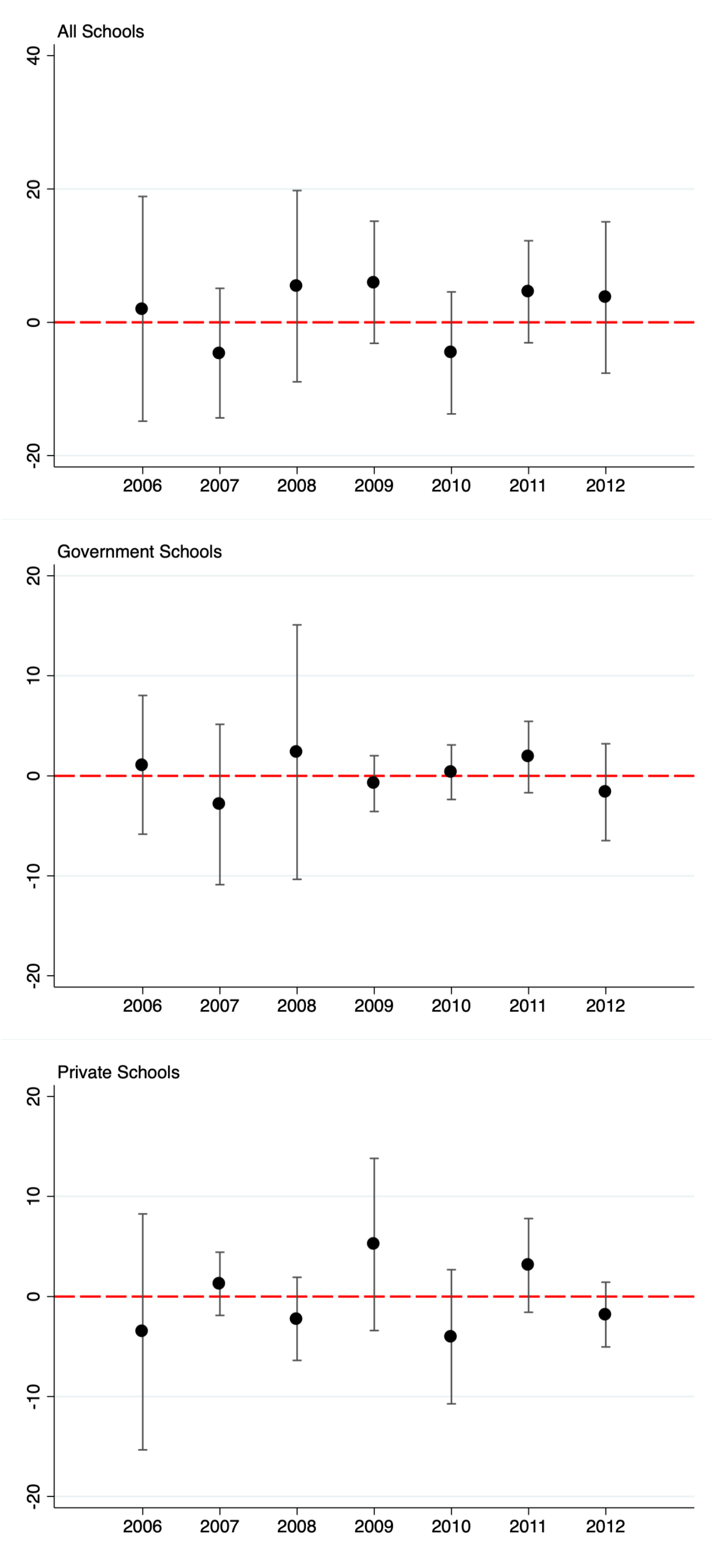

Figure A2: **Discontinuity in the annual number of schools constructed in post-DPEP period**

Notes: These graphs plot the (district-level) RD estimate of the number of schools constructed each year. The results suggest that there are no significant differences in the number of schools constructed across the RD cutoff in the post–treatment years (2006 to 2012). The standard errors are robust-bias corrected and clustered at the district level.

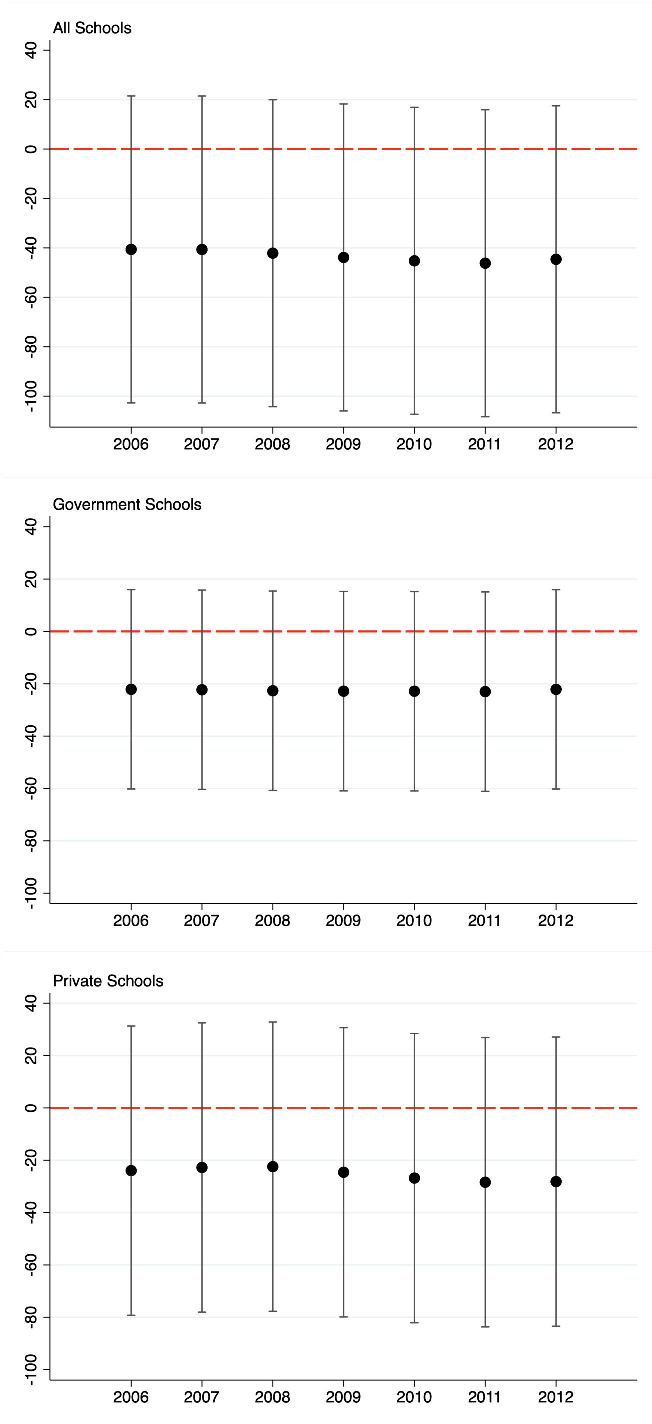

Figure A3: **Discontinuity in the total number of schools in post-DPEP period**

Notes: These graphs plot the (district-level) RD estimate of the total number of schools present in any given year. The results suggest that there are no significant differences in the number of schools across the RD cutoff in the post–treatment years (2006 to 2012). The standard errors are robust-bias corrected and clustered at the district level.

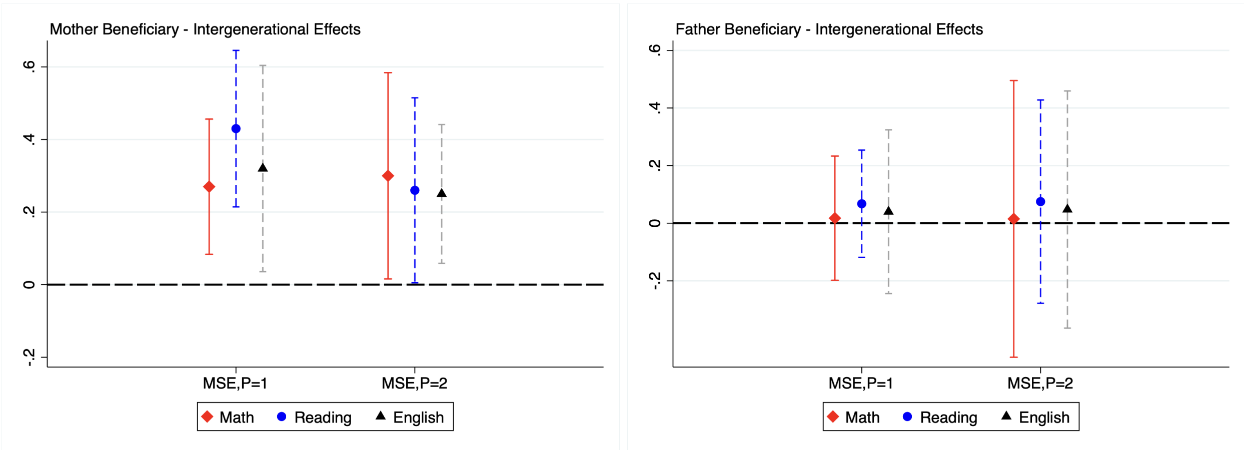

Figure A4: **Robustness Check – optimal bandwidth using MSE method**

Notes: The sample consists of children who satisfy two conditions - (1) their mother was born between 1980 and 1995, and (2) they started school after 2005 (end of DPEP). All specifications control for categorical variables for gender of the child, state and year of data collection. The standard errors are clustered at the district level. Data Source: Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018.

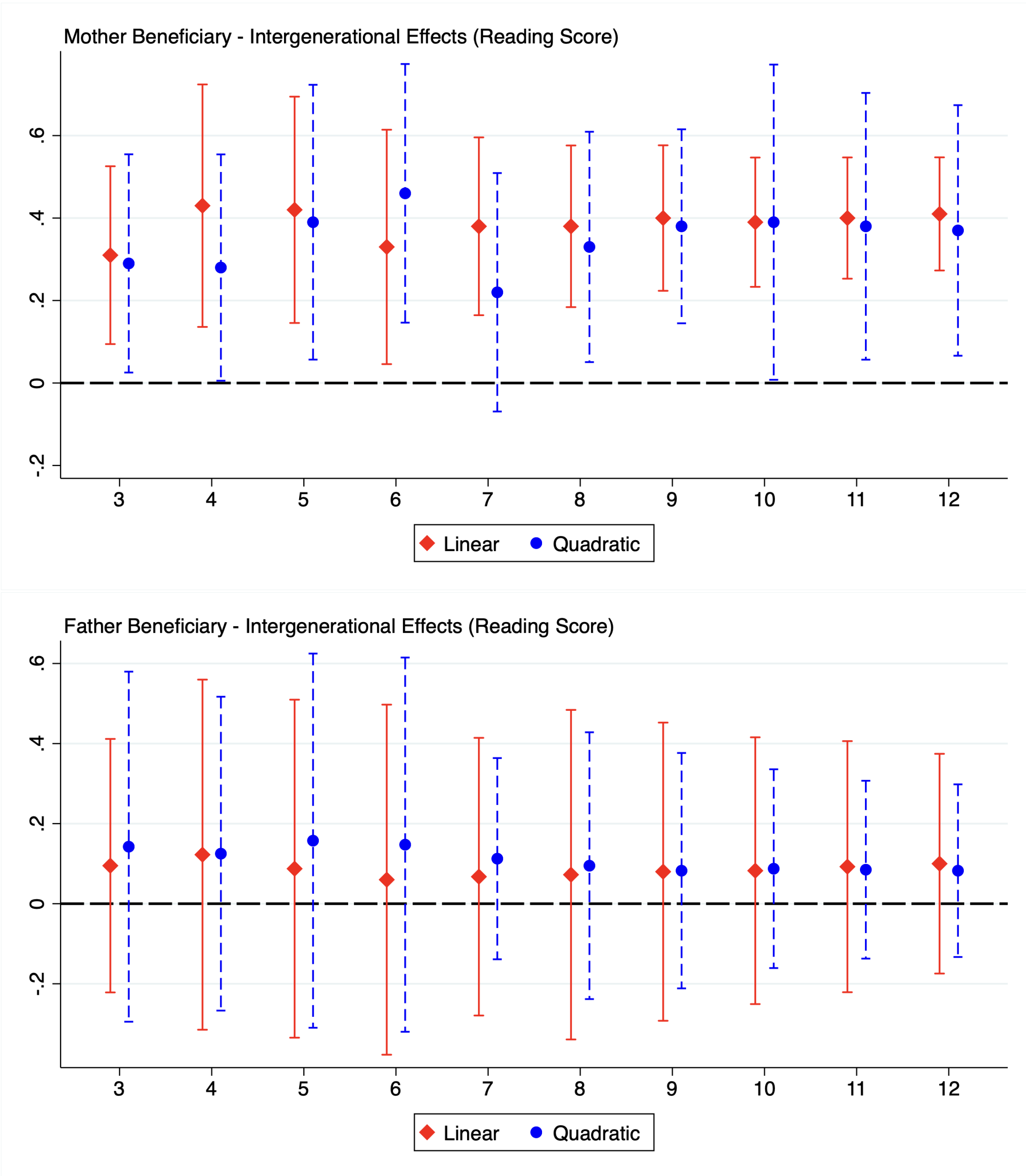

Figure A5: **Robustness to different bandwidths and polynomial functions – Reading**

Notes: The sample consists of children who satisfy two conditions - (1) their mother/father was born between 1980 and 1995, and (2) they started school after 2005 (end of DPEP). All specifications control for categorical variables for gender of the child, state and year of data collection. The standard errors are clustered at the district level. Data Source: Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018.

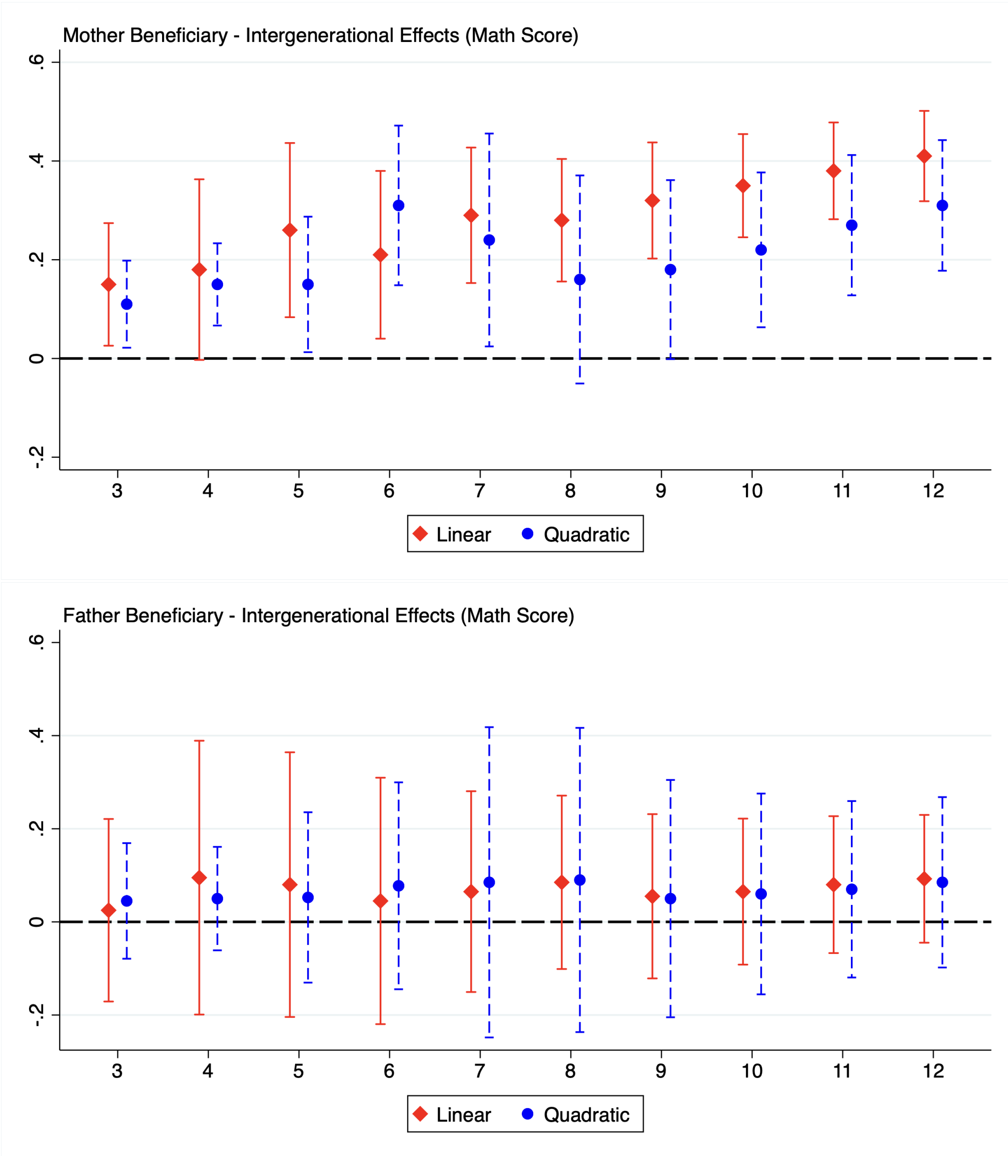

Figure A6: **Robustness to different bandwidths and polynomial functions – Math**

Notes: The sample consists of children who satisfy two conditions - (1) their mother/father was born between 1980 and 1995, and (2) they started school after 2005 (end of DPEP). All specifications control for categorical variables for gender of the child, state and year of data collection. The standard errors are clustered at the district level. Data Source: Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018.

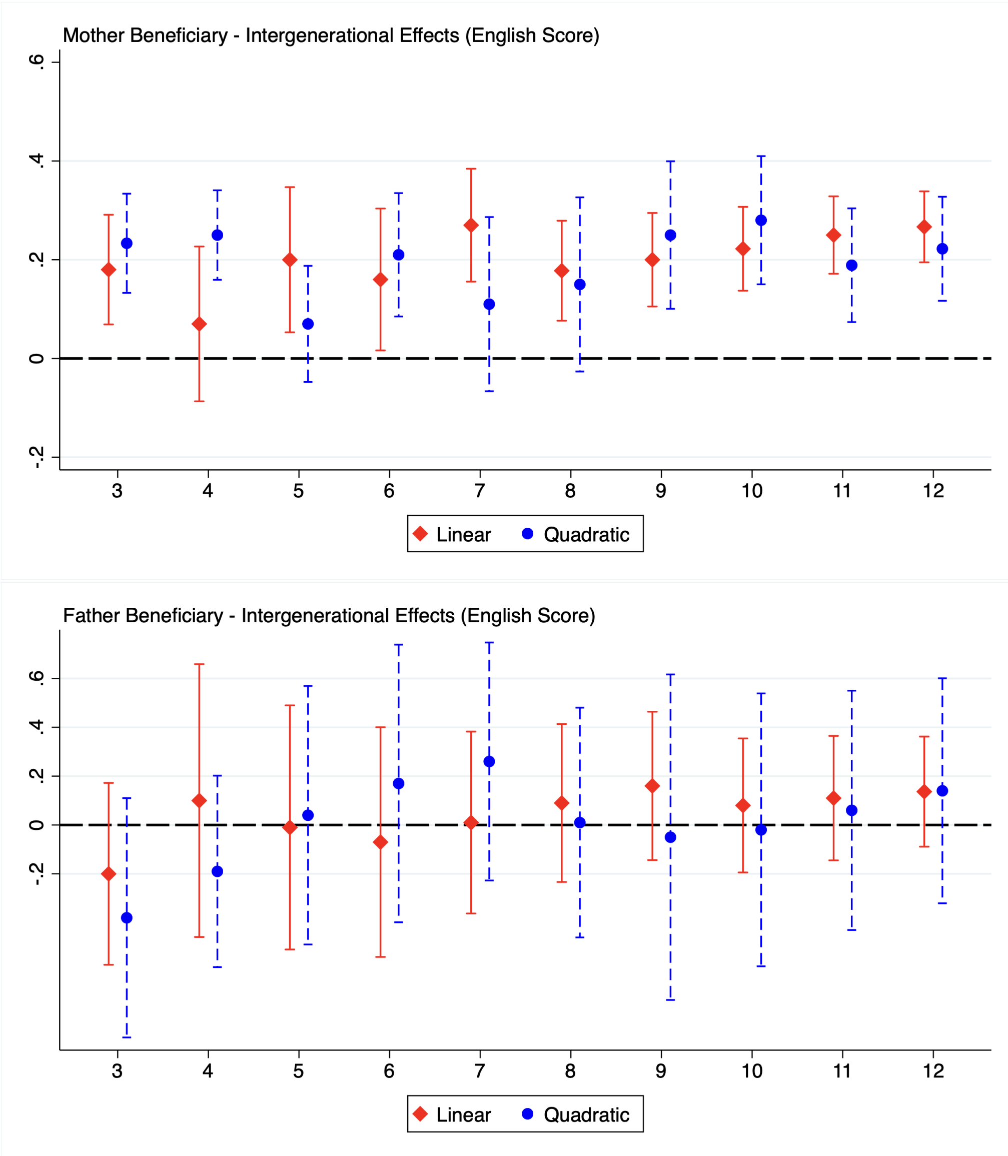

Figure A7: **Robustness to different bandwidths and polynomial functions – English**

Notes: The sample consists of children who satisfy two conditions - (1) their mother/father was born between 1980 and 1995, and (2) they started school after 2005 (end of DPEP). All specifications control for categorical variables for gender of the child, state and year of data collection. The standard errors are clustered at the district level. Data Source: Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018.

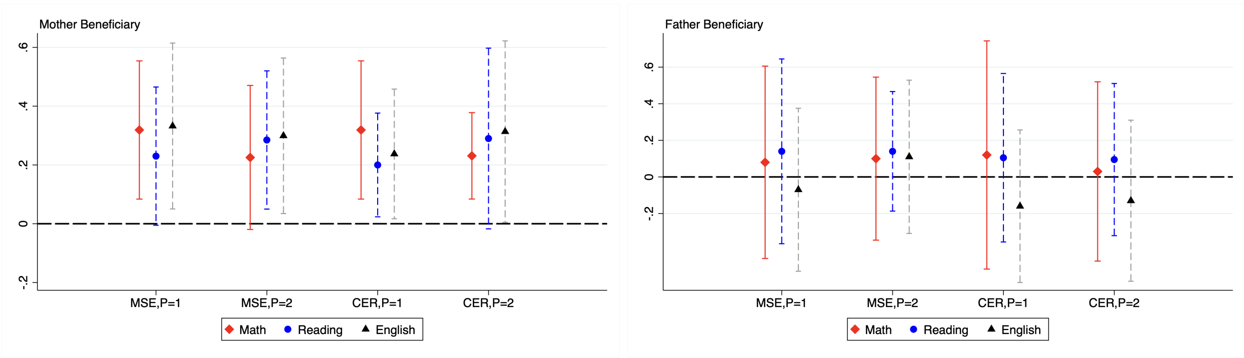

Figure A8: **Robustness Check – using epanechnikov kernel**

Notes: The sample for this analysis consists of children who satisfy two conditions - (1) their mother/father was born between 1980 and 1995, and (2) they started school after 2005 (end of DPEP). The RD point estimates are constructed using the epanechnikov kernel. All specifications control for categorical variables for gender of the child, state and year of data collection. The standard errors are clustered at the district level. Data Source: Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018.

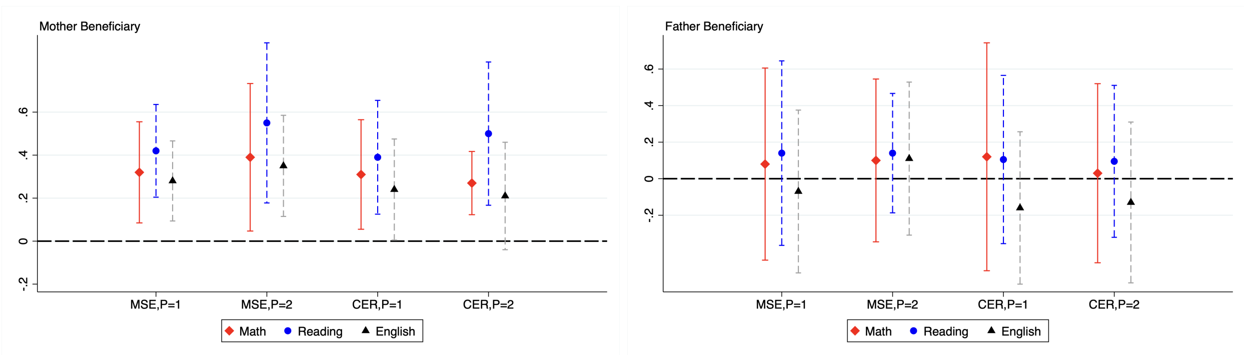

Figure A9: **Robustness Check – considering 5–12 year olds as direct beneficiaries**

Notes: The sample for this analysis consists of children who satisfy two conditions - (1) their mother/father was born between 1982 and 1995, and (2) they started school after 2005 (end of DPEP). All specifications control for categorical variables for gender of the child, state and year of data collection. The standard errors are clustered at the district level. Data Source: Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018.

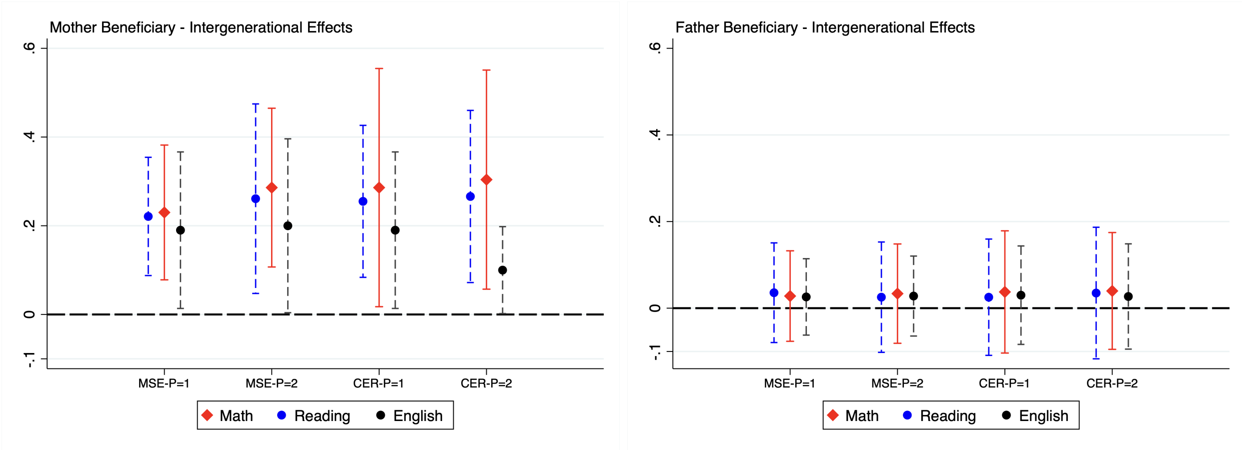

Figure A10: **Robustness Check – DPEP star using archive data**

Notes: The sample for this analysis consists of children who satisfy two conditions - (1) their mother/father was born between 1980 and 1995, and (2) they started school after 2005 (end of DPEP). All specifications control for categorical variables for gender of the child, state and year of data collection. The standard errors are clustered at the district level. Data Source: Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018. The start year of DPEP in treatment regions comes from archival data (section 5.4.1) and the control regions use 1993 to 2000 as the reference period.

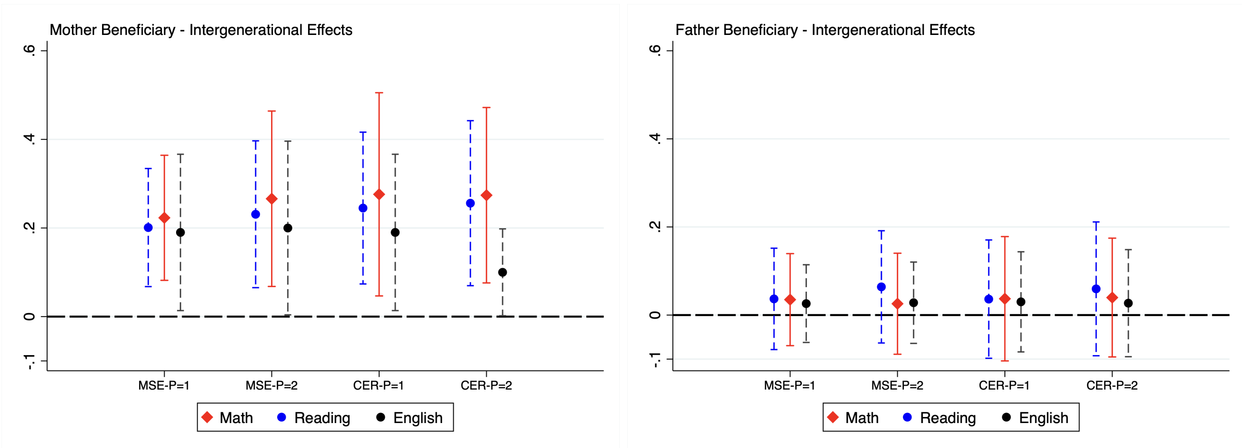

Figure A11: **Robustness Check – Archival start date (variation)**

Notes: The sample for this analysis consists of children who satisfy two conditions - (1) their mother/father was born between 1980 and 1995, and (2) they started school after 2005 (end of DPEP). All specifications control for categorical variables for gender of the child, state and year of data collection. The standard errors are clustered at the district level. Data Source: Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018. The start year of DPEP in treatment regions comes from archival data (section 5.4.1) and the control regions use state-wise average start date of DPEP to assign DPEP years.

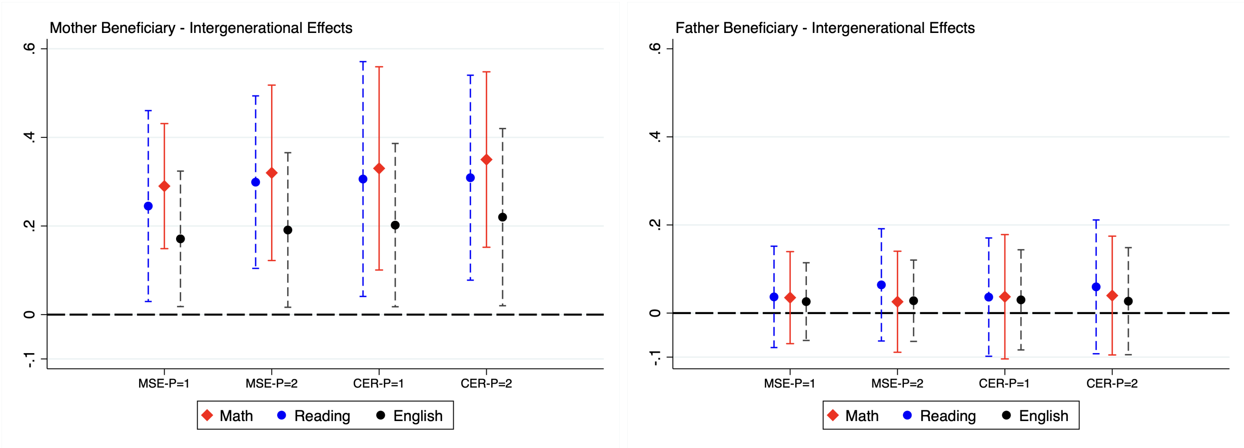

Figure A12: **Robustness Check – DISE start time**

Notes: The sample for this analysis consists of children who satisfy two conditions - (1) their mother/father was born between 1980 and 1995, and (2) they started school after 2005 (end of DPEP). All specifications control for categorical variables for gender of the child, state and year of data collection. The standard errors are clustered at the district level. Data Source: Annual Survey of Education Report (ASER) individual level data for the years 2007-2014, 2016 and 2018.

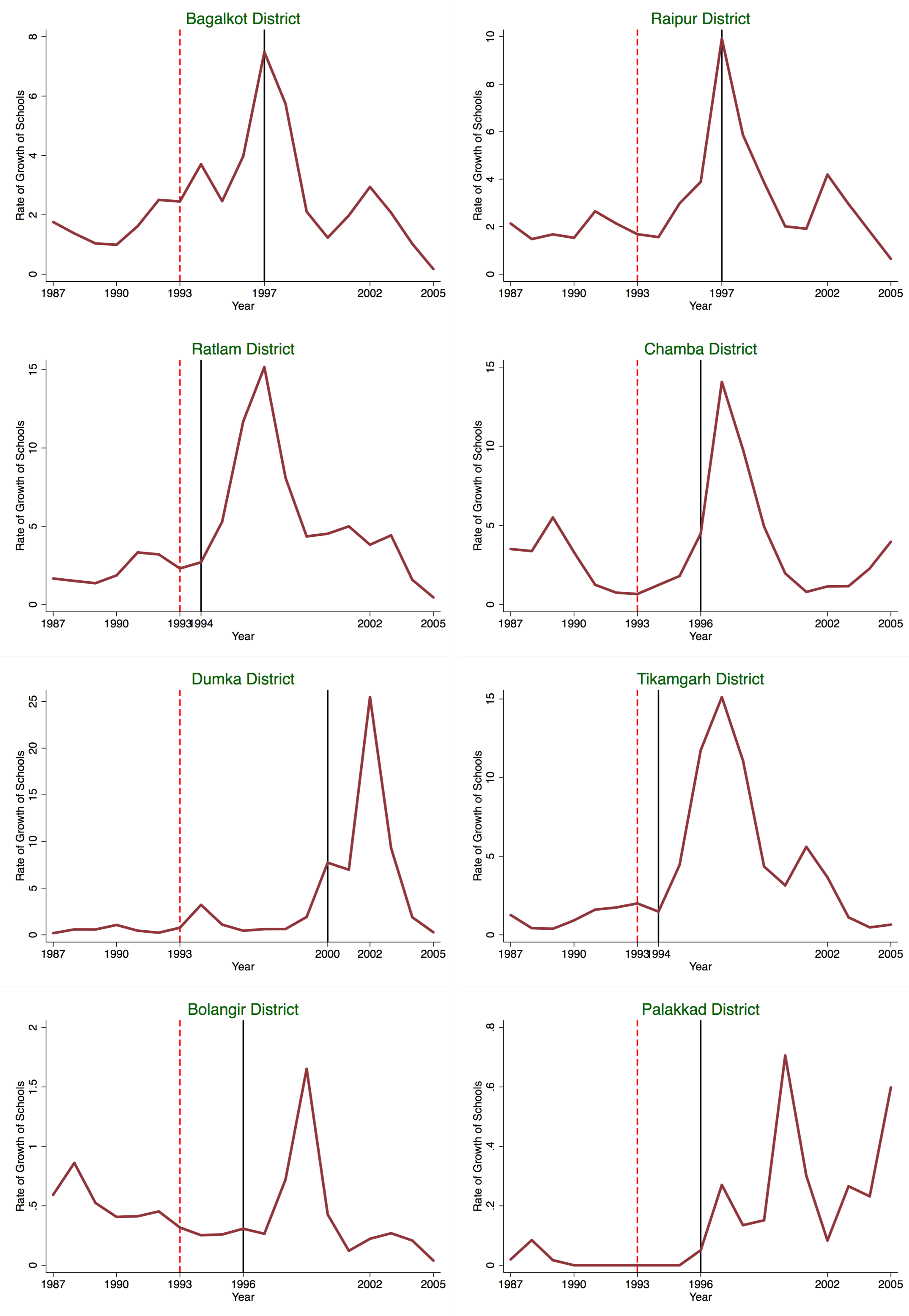

Figure A13: **Yearly rate of growth of school construction - Treatment Districts**

Notes: The graphs in this figure plot the annual rate of growth of schools in a particular district. The dashed (red) line shows the year when the DPEP policy was announced (1993), the bold (black) line represents the year of implementation inferred from the archival documents (*Archival start year*) and the peak of the rate of growth distribution is an alternative start date (*DISE start year*) of the policy. Based on DISE 2014 data.

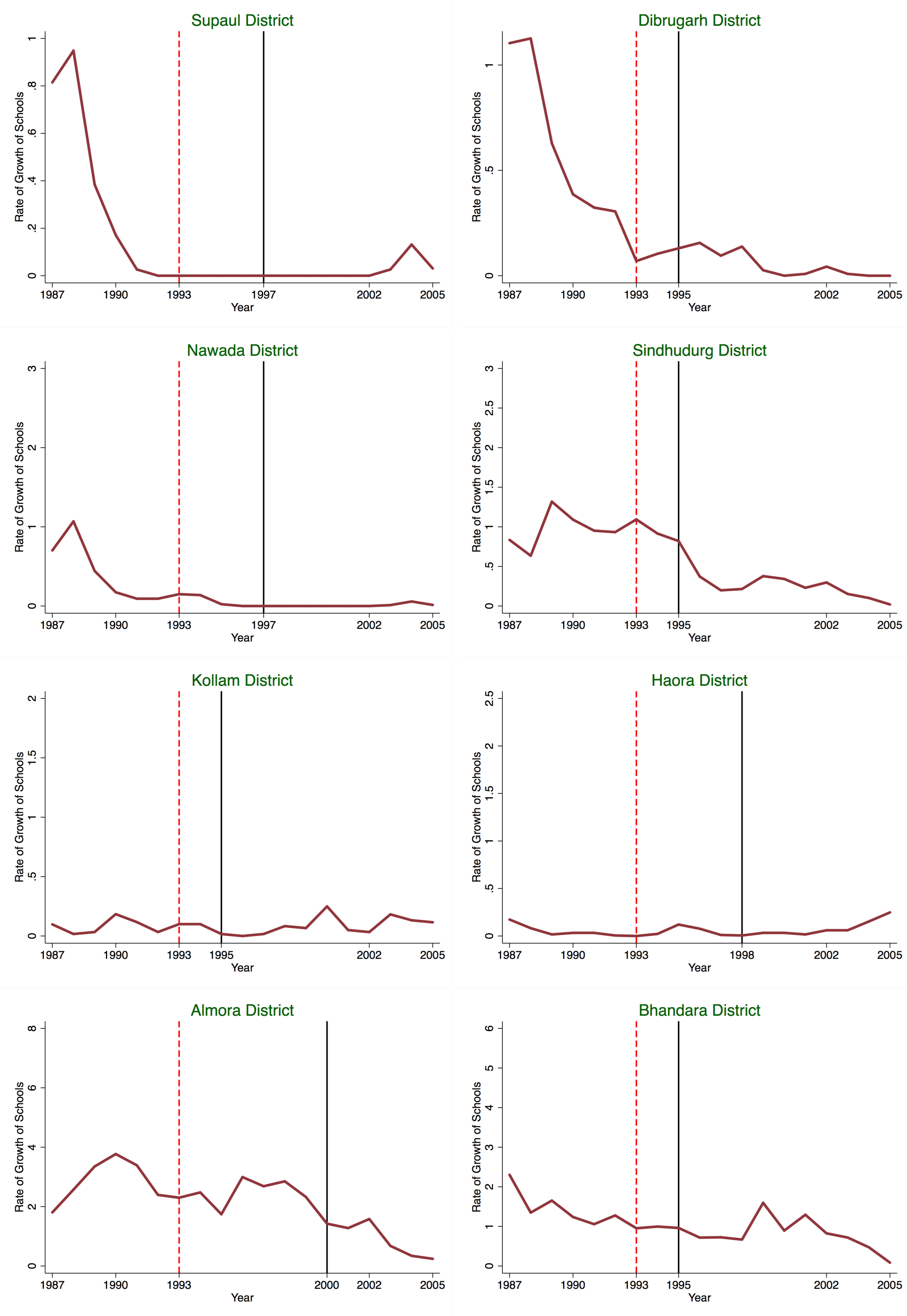

Figure A14: **Yearly rate of growth of school construction plotted against time - Control Districts**

Notes: The graphs in this figure plot the annual rate of growth of schools in a particular district. The dashed (red) line shows the year when the DPEP policy was announced (1993), the bold (black) line represents the year of implementation inferred from the archival documents (*Archival start year*). Based on DISE 2014 data.

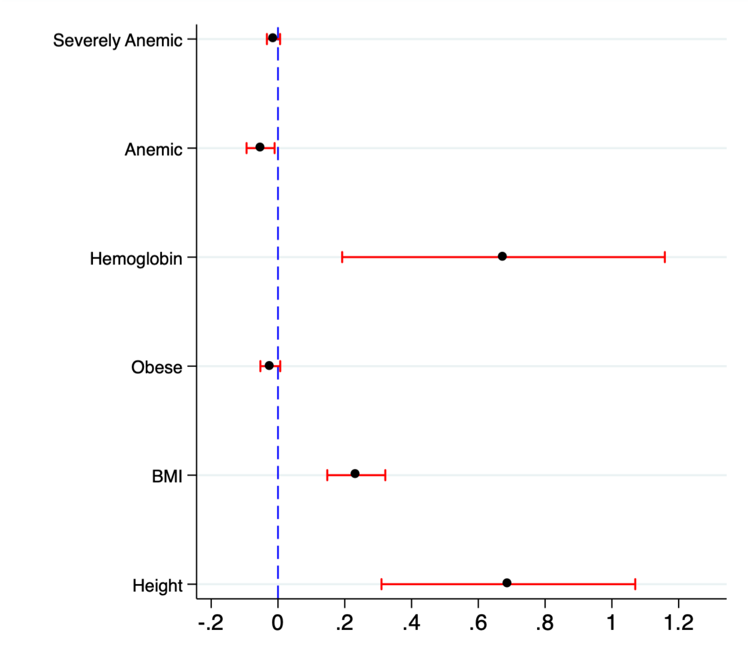

Figure A15: **Mechanisms - Health of direct beneficiary women**

Notes: The sample consists of women who were below 14 years of age at the time of implementation of the DPEP programme, that is they were born between 1980 and 1995. All the specifications control for categorical variables for caste, religion and state. The standard errors are robust-bias corrected and are clustered at the district level. Data Source: NFHS 2015-16.

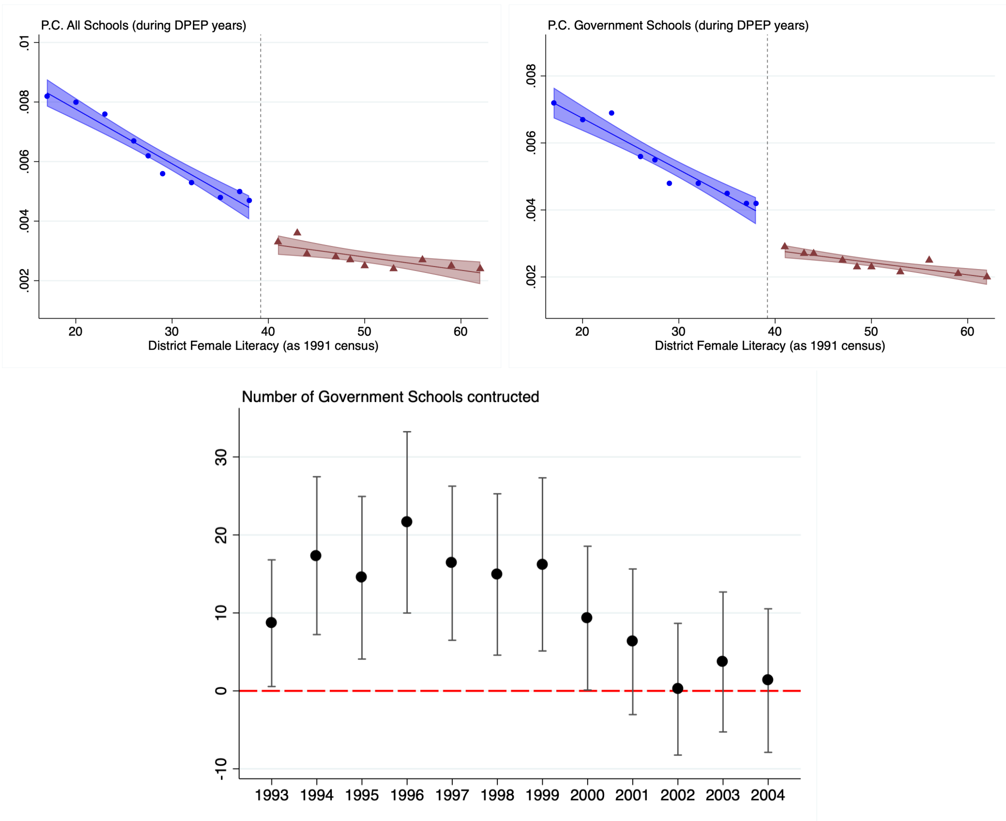

Figure A16: **Quantity of schooling (during DPEP years)**

Notes: Upper panel: Total number of schools that were constructed in the pre-DPEP period (before 1993). Bottom panel: Per capita schools constructed during the programme years. Bottom Panel: Year-wise RD coefficient for the total number of government school constructed annually (outcome variable).

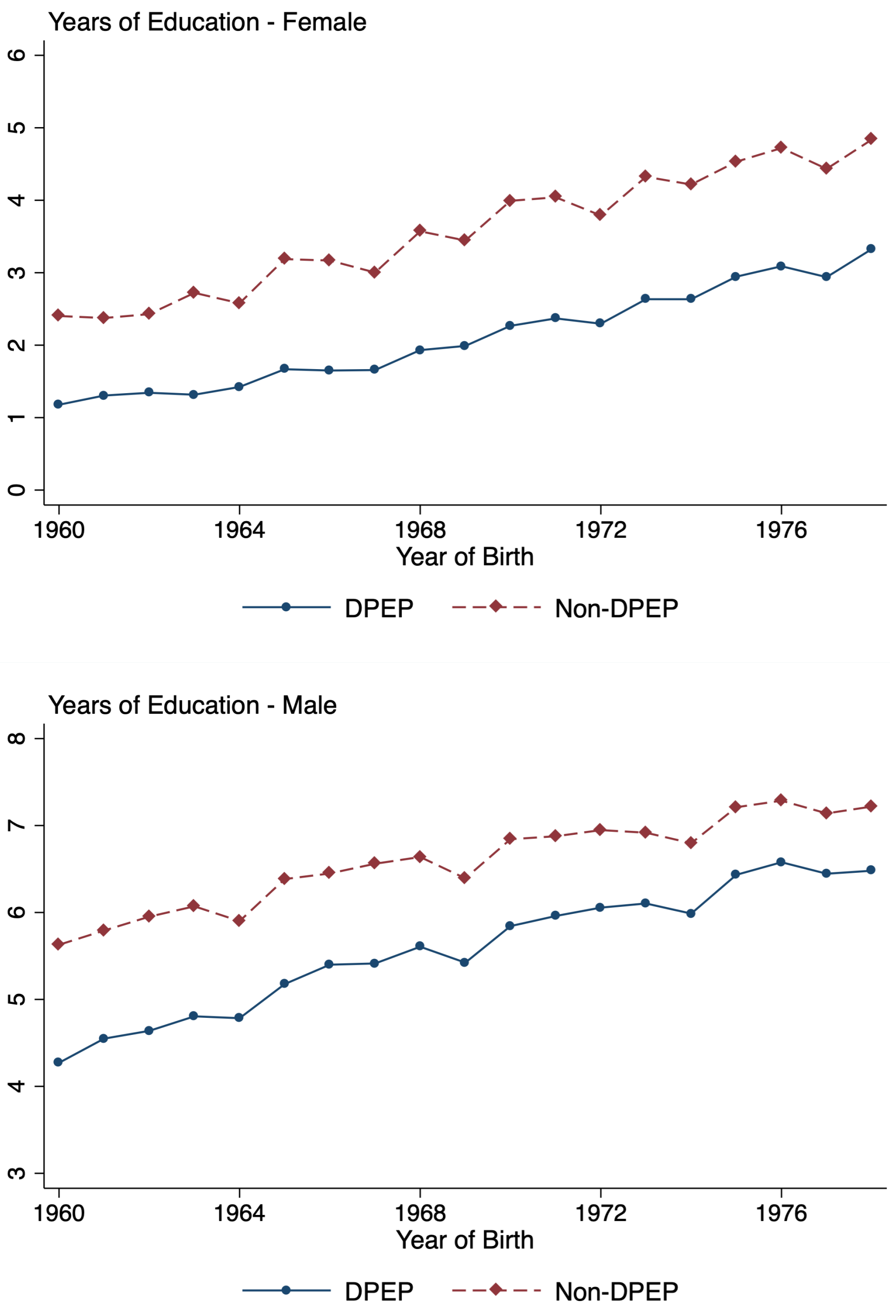

Figure A17: **Pre-trends in the raw data**

Notes: The figure in the top panel plots the mean education for women in each birth-year (between 1960 and 1978) for the DPEP and non-DPEP districts separately. The bottom panel does the same for men. Individuals born between 1960 and 1978 would have been unaffected by DPEP since they were 16 to 34 years of age at the start of the DPEP policy in 1993-94. Therefore, this figure shows that there were parallel trends in years of education between the treated and control districts. Data: DLHS Rounds 3 and 4.

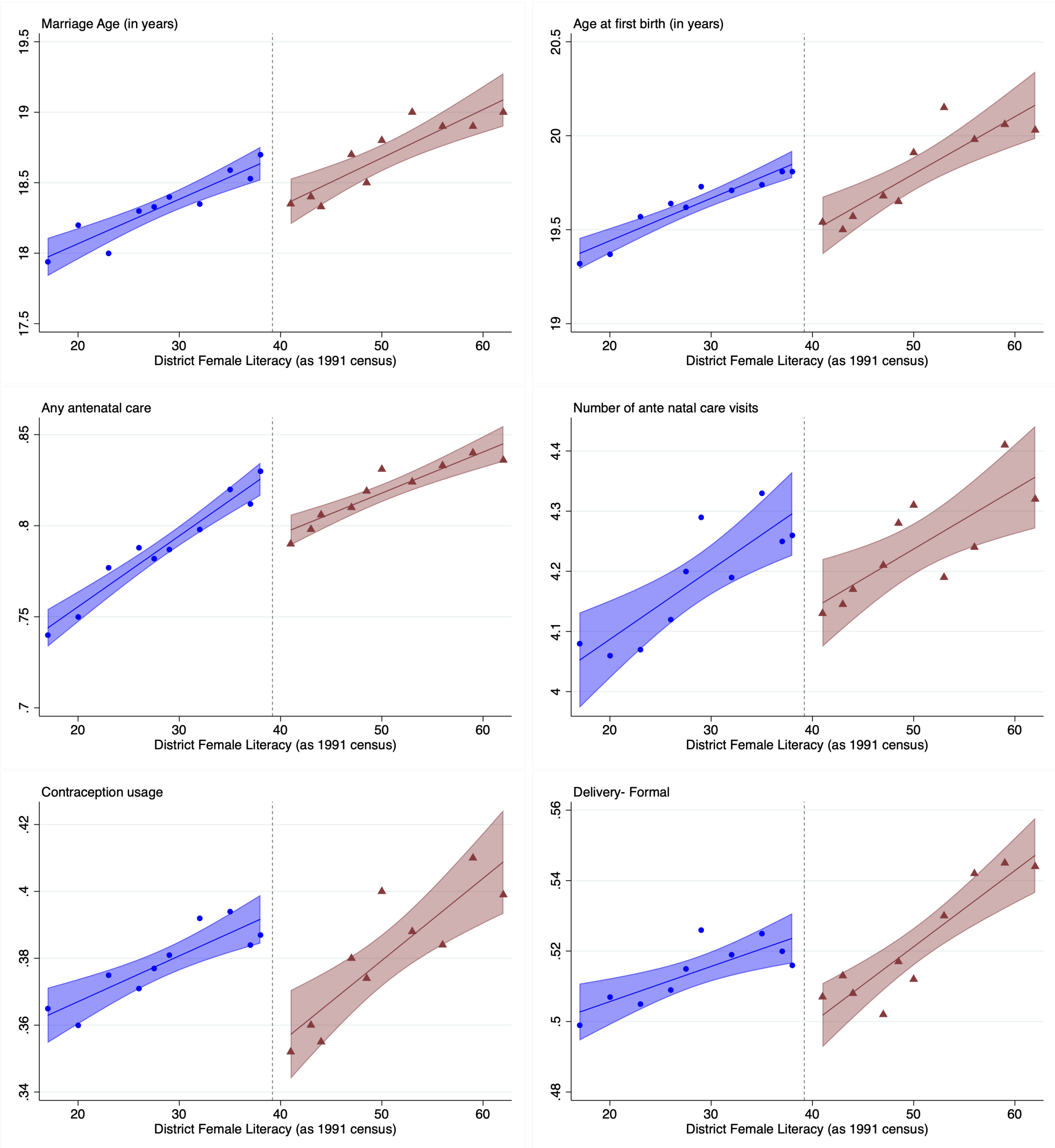

Figure A18: **Effect of DPEP on women's outcomes – potential mechanisms**

Notes: The sample consists of women who were born between 1980 and 1995. Marriage age refers to the age at marriage (in years), Age at first birth refers to age when the woman had her first child (in years), Contraceptive use is a dummy that takes a value of one if the women reported using contraceptives, *Any ANC* is a categorical variable that takes a value of one if the woman accessed any ANC facilities during the last pregnancy, # ANC visits refers to the number of ANC visits made during the last pregnancy, *IFA Taken* is a categorical variable that takes a value of one if the woman reported taking Iron and Folic Acid (IFA) tablets during the last pregnancy, *Delivery-Formal* is a categorical that takes a value of one if the woman reported giving birth in a formal health facility and *Any PNC* refers to a dummy that takes a value of one if the woman reported using any Post Natal Care (PNC) facilities. Data Source: District Level Household and Facility Survey (DLHS) data Round 3 (2007-08).