DISCUSSION PAPER SERIES

IZA DP No. 5274

Can Private School Growth Foster Universal Literacy? Panel Evidence from Indian Districts

Sarmistha Pal
Geeta Gandhi Kingdon

October 2010

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Can Private School Growth Foster Universal Literacy? Panel Evidence from Indian Districts

Sarmistha Pal
Brunel University
and IZA

Geeta Gandhi Kingdon
IoE, University of London,
CSAE Oxford and IZA

Discussion Paper No. 5274
October 2010

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 5274
October 2010

ABSTRACT

Can Private School Growth Foster Universal Literacy? Panel Evidence from Indian Districts*

Millennium Development Goals (MDGs) set the agenda for the attainment of universal literacy by 2015 primarily to be delivered by the state sector. This agenda tends to ignore the significant private school growth around the world since early 1990s, thus initiating the policy debate as to whether private school growth may foster 'education for all'. Despite growing literature on the difficulties of attaining MDGs, there is hardly any attempt to assess the role of private sector in this respect. Using India as an important case in point, we intend to bridge this gap of the literature. Results using a unique district-level panel data-set from 17 major states of India for the period 1992-2002 that we compile highlight a significant positive impact of private school growth on literacy while its effect on gender gap in literacy remains rather limited in our sample. Compared to 15-19 year olds, private school effect of literacy is stronger among 10-14 year old children. Interesting variations across the regions and also among the marginalised ethnic groups are noted. The paper offers explanations for the findings.

JEL Classification: I21, I28, O15

Keywords: private school growth, universal literacy, gender gap, Millennium Development Goals, India, Asia

Corresponding author:

Sarmistha Pal
CEDI, Department of Economics and Finance
Brunel University
Uxbridge UB8 3PH
United Kingdom
E-mail: sarmistha.pal@brunel.ac.uk

* Sarmistha Pal wishes to thank Leverhulme Trust for funding this research, NCERT for access to AISES data and CSAE, Oxford University for their hospitality where much of this research was done. We would also like to thank Petia Topalova for poverty data and Mohammed Arif, Ana Maria Lugo, Yama Temouri for excellent research assistance, Paul Atherton, Erlene Berg, Sonia Bhalotra, Irma Clots-Figures, Jishnu Das, Aparajit Mahajan, Pushkar Maitra, Karthik Muralidharan and also participants at the UKFIET Conference on Education and Development at Oxford, Economic Growth and Development Conference at the Indian Statistical Institute New Delhi, and Social Policy in India workshop at Warwick University for many constructive comments/suggestions. Any remaining errors are ours.

Can Private School Growth Foster Universal Literacy? Panel Evidence from Indian Districts

1. INTRODUCTION

In April 2000 the World Education Forum's meeting in Senegal set the agenda for the attainment of universal primary education by 2015. It noted that the 'prime responsibility' for achieving universal education lay with national governments, assisted by aid agencies. This prescription did not however take account of the burgeoning private schools in different parts of the world including some emerging economies - an important oversight which subsequently led to a policy debate as to whether growth of private schools could be conducive to universalising literacy (Tooley, 2004; Watkins 2004). While there have been some assessments about the prospects of meeting the Millennium Development Goals (MDGs) by 2015 around the world (Bloom et al. 2006), we are not aware of any analysis assessing the potential role of private provision of education on universal literacy. The present paper aims to bridge this gap of the literature.

It has often been argued that greater market orientation makes private schools and teachers more accountable to parents, more sensitive to input costs and thereby more efficient. A fast pace of private school growth could however raise concerns for equity reasons, especially in the light of the MDGs. First, private schools are the fee paying schools and hence would naturally exclude children from poorer/disadvantaged background. Further, given the importance of son preference especially in some Asian countries, private school growth could widen the gender gap between boys and girls if this induces resource constrained parents to send only their boys to private schools, thus encouraging discrimination against girls. Given this trade-off, it is pertinent to explore whether and how the efficiency argument in favour of private school growth could be compatible with the equity argument, directly linked to MDGs of securing universal literacy by 2015.

India is an important case in point. While the state sector still dominates the schooling market in India, an important feature of the 1990s has been a significant growth of private schools in India (PROBE 1999). While there is a large and growing literature on child schooling in developing countries (e.g., see Glewwe, 2002; Hanushek and Woessman 2008), there is very little direct evidence, if at all, about the possible effects of recent growth of private schools on universal literacy. Most available studies highlighted the relative efficiency of private schools in imparting learning, using information from small scale surveys in particular regions of India. For example, Kingdon (1996) found that students in private schools performed significantly better than those in government schools in urban Lucknow district, after controlling for student background and sample selectivity. Bashir (1994) indicated that students in private schools had better Mathematics achievement, but less achievement in Tamil language, compared to government school students in Tamil Nadu.[1] Using recent national IHDS and ASER data from major Indian states, Desai et al (2008) and French and Kingdon (2010) respectively found that among children from the same household (i.e., who are likely to share similar characteristics in terms of ambition, ability and motivation, and in terms of the household environment), those who attend private school performed better in language and maths than those who attend government school. Two further studies have analysed the factors behind private school growth at a broader level. Considering a nationally representative sample from major Indian states, Muralidharan and Kremer (2008) argued that private schools are more likely to be set up in areas where state schools are failing. Using PROBE data from five north Indian states, Pal (2010) found that private schools are more likely to be present in villages with better off households and better infrastructural facilities, while the effect of private school growth on government school pass rates remains insignificant.

So the question remains whether and how private school growth can foster universal literacy. Increase in private school share at a given level could affect literacy

[1]Many PUA schools emphasise English or specialise in English as a medium of instruction, and thus may have fewer hours devoted to Tamil teaching. There is however no consensus in the literature about the effect of school type on school quality. For example, for Indonesia Newhouse and Beegle (2005) found that students from public secondary schools perform better than comparable privately schooled students but Bedi and Garg (2000) argued that graduates of private secondary schools perform better in the labour market in the same country.

through affecting demand and/or supply. On the one hand, private schools are directly accountable to parents and children and thus tend to resolve the incentive problems commonly present in the management of government schools (incentive effect).[2] On the other hand, the apparently greater effectiveness of private schools may be linked to the unobserved characteristics of parents/students attending private schools (selection effect). The pertinent issue here is to explore whether an increased share of private unaided schools could boost literacy for all in the aggregate (at the district level). If it did not, private school growth, in a sense, fails to be compatible with the objective of attaining universal literacy by 2015, which is a corollary of the education related MDG.

In principle, a higher private school share could either raise or lower literacy rates, or indeed have no effect on literacy. It may raise literacy if private schools impart higher learning than government schools, as has been found by some micro-level studies. It may also raise the literacy if presence of private schools boosts the quality of local government schools (e.g., Hoxby, 1994), which is not supported by existing evidence (Pal 2010). It may lower literacy rates if the growth of private schools causes the closure of or deterioration in the quality of government schools. Private school growth may have no impact on literacy as such; other things remaining unchanged, if, for example, those who choose private schooling are relatively better-off and/or more motivated towards schooling and would have become literate even in the absence of private schooling (e.g. via enrolment in government schools with/without private tuition), privatisation would not have a significant impact on literacy rates.

A related issue is the implication of private school growth for gender differences in literacy rates. This is particularly important for a country like India where a pronounced gender difference in literacy persists, especially in the large north Indian states such as Bihar, UP, MP and Rajasthan. Private school growth could potentially exacerbate gender difference in literacy rates since it can enable parents characterised by pronounced son-preference to exercise that preference by sending sons to private and daughters to

[2] Another possible channel would be to assess the effect of school privatisation on performance of government schools. We however do not have any information on performance indicators of students attending private and public schools; hence we were unable to test this hypothesis. We also note that Pal (2008) argued that presence of local private schools fails to have a significant impact on government school pass rates in the PROBE villages.

government schools.[3] If learning levels are better in private schools, the gender gap in learning (and literacy) may increase with private school growth as girls, especially in the presence of son preference in school choice, are likely to be excluded from private schools. On the other hand, private schools may mitigate gender differences in educational outcomes. For example, if private schools fulfil differentiated demand (e.g., provide local schools so that girls do not have to travel far or provide separate toilets for girls and boys), availability of private schools will increase girls' access to schooling and learning and thus reduce the gender gap in literacy. However, it remains possible that any apparently beneficial effect of private schools (in terms of a reduced gender gap in schooling/literacy) is in fact due to an aspect of sample selection. Parents who choose to send daughters to private schools (especially at the higher levels of education) are not a random draw from the population of all parents; they are likely to be the more enlightened in terms of attitudes to gender equality. Thus, if a higher share of private schooling in a district is associated with a lower gender gap in literacy, this could potentially be a spurious relationship, simply due to sample selection, i.e. in reality there may be no relationship between the two, Thus, on balance, whether the growth of private schooling has a negative, positive or neutral association with gender gap in literacy rates, ceteris paribus, remains an open empirical question that we explore here.

Studies that analyse different aspects of private schooling growth in India have primarily used single-year cross-section data. Consequently, existing estimates of the effects of private schooling are likely to suffer from endogeneity bias primarily due to unobserved heterogeneity among market participants (schools/parents/children). For example, in an achievement production function, a private school dummy variable is endogenous since it is likely to pick up the effect of child or family level unobserved factors (e.g. motivation, ambition etc.) that make it more likely that a child will attend private school and also raise achievement levels. Similarly, to extend to district level data, if we were to regress district literacy rate in a given year on the district's 'private school share' in that year, the latter variable would suffer from omitted variable bias in such an

[3] There is however no evidence that parental wealth is associated with greater gender difference in literacy in India (e.g., see Pal, 2004).

OLS regression. There could also be reverse causality: just as private school presence affect achievement/literacy, the latter may also influence private school presence.

We use a unique two-period district-level data for 1992 and 2002 from 17 major Indian states compiled from various official sources. Access to a this district-level panel data allows us to address the underlying econometric issues while examining the effect of private school growth on literacy rates (a measure of educational achievement) and the associated gender gap in literacy, two of the key millennium development goals (MDGs). The extent of private schooling in a district is measured by the share of private schools[4] (at a given level, namely, upper primary and secondary) in total schools at that level.[5] We start with a district fixed effects approach (which is equivalent to the first difference estimates as we have data only for two periods) that allows us to obtain consistent estimates, net of the effect of any time-invariant district-specific unobserved characteristics. Further we need to minimize the potential bias generated by reverse causality from literacy to private school growth. While randomized experiments provide a good solution to the problem of endogeneity in general (e.g., see Duflo and Hanna, 2005), they cannot be used here since private school presence cannot be randomly allocated. Accordingly, we compare FE-OLS estimates with ols estimates of changes in literacy as well as that in gender gap in literacy over 1992-2002 (in terms of 1992 levels of private schools share and other explanatory variables) and also fixed-effects instrumental variable (FE-IV) estimates with a view to redress biases generated by both simultaneity and unobserved heterogeneity. Note however that the identification of a convincing instrumental variable is not straightforward (see further discussion in section 3).

We focus our attention on children aged 10-19 years old and classify them into two subgroups: upper primary school aged children (10-14 years old) and the secondary school aged children (15-19 years old). For a given age-group, we also compare the full sample estimates with estimates from various sub-samples: (a) districts located in large north (as opposed to southern districts) Indian states who are generally less well-off and known for greater discrimination against female children; this regional comparisons allow us to

[4] See section 2 for a discussion of types of private schools in India.

[5] Henceforth we use private school growth and growing private school share at a given level of schooling interchangeably for the rest of this paper.

explore whether the central results hold in rather diverse regions; (b) districts targeted by the District Primary Education Programme (DPEP), which was a programme to boost literacy in worse performing regions (see further discussion in section 4.2.2). (c) sub-sample of SC/ST children within districts. Since information on age-specific SC/ST literacy rates are available only for 2001 Census, we are forced to use single cross-section data to analyse the effect of private school share on SC/ST literacy rates for 10-19 year olds. [6]

Results suggest that estimated coefficients of private school share are bigger in FE-IV model relative to FE-OLS model. In other words, FE-OLS estimates tend to under-estimate the private school effect of literacy. This seems convincing because districts with higher literacy are generally the better-off states which tend to have lower share of private schools (e.g., see Muralidharan and Kremer, 2006; Pal, 2010), thus generating a negative simultaneity bias. Direction of the relationship is however independent of the choice of regression models. In general, there is evidence of a significant positive impact of private school growth on literacy though its effect on gender gap in literacy is rather weak in our sample. Compared to 15-19 year olds, literacy effect of private school share is stronger for 10-14 year olds, which is perhaps convincing: 10-14 year olds in 2002 were the first group of children to enjoy the full benefits of pronounced private school growth that sets a trend from early to mid 1990s. Further evidence suggests that there are important regional variations not only between the northern and the southern states, but also between districts that opted for District Primary Education Programme (DPEP) and those who did not. Finally, private school growth has significant literacy effect on all SC children but only 10-14 year old ST children and the effect is stronger for ST children.

The paper is developed as follows. Section 2 describes the data while section 3 explains the methodology. Results are discussed in section 4. The final section concludes.

2. DATA

[6] Note however that the available data do not allow us to distinguish between *gender gap* in literacy for SC/ST children.

Data has been compiled from various sources: This includes the Sixth (1992-93) and Seventh (2002-03) All India School Education Survey (AISES) data and also Census data (1991 and 2001). District-level AISES data cover information on the number of *recognised* schools (by management type, i.e., private/public, etc.), enrolment by gender and caste (scheduled castes, SC; scheduled tribes, ST), characteristics of teachers (gender/caste), and physical facilities at primary, upper primary and secondary levels of schooling in the district. District level census data from 1991 and 2001 provide information on population composition (by gender/caste); literacy rates for different age categories of the population (male/female and total); and access to various infrastructural facilities. In addition, we obtain district-level poverty head count rates information from the 50th (1993-94) and 55th (1999-00) rounds National Sample Survey (NSS) data. We merge 1991 Census data and 50th round NSS data with 6th AISES to generate district-level information for 1992. Similarly, we merge 2001 Census data and 55th round NSS data with 7th AISES data to generate the corresponding district-level information for 2002. This allows us to build up a two-period panel data for the period 1992-2002.

There are three broad types of *recognised* schools in India, namely, government schools, private aided schools (PA) and private unaided schools (PUA) schools.[7] Government and aided schools are invariably 'government-recognised', i.e. they have the government stamp of approval. They are similar to each other in many respects since aided schools are almost entirely financed by the government and have little control over staffing (hiring/firing) and fee levels, despite being nominally privately managed.[8] PUA schools (whether recognised or not) are more autonomous than aided schools and are totally self-funded out of fee income. Thus PUA schools are the truly 'private' schools in India. At the secondary school level, all schools have to be government-recognised. But at the primary

[7] In order to receive recognition, however PA and PUA schools must fulfil several requirements that are prohibitively expensive for many schools, especially those serving the poor (e.g., hold a sizeable cash bond with the government, provide sizeable playgrounds, etc.).

[8] There is some inter-state variation in the management of PA schools. For example, PA schools in Uttar Pradesh have no control over hiring/firing of own teachers (who are appointed by the UP School Service Commission). In contrast, PA schools in Tamil Nadu have some autonomy to select and hire their own teachers.

school level, many PUA schools remain unrecognised.[9] Non-recognised schools are not included in any government list of schools and are thus not included in the periodic school 'census' (called the 'All India School Education Survey' (AISES). As a result, our analysis in this paper can only include the *recognised* PUA schools rather than all PUA schools. This is an unfortunate but an inevitable data limitation since there is no source that provides information on unrecognised PUA schools for all districts of India going back to early 1990s[10]. However, in general, there is likely to be a positive correlation between the share of recognised PUA schools and the share of all PUA schools (recognised and unrecognised) since districts that have more recognised PUA schools are also likely to be the districts that have more unrecognised PUA schools. As such our result would provide only a lower bound of the growth of private schools in the Indian districts. In the rest of the paper whenever we refer to PUA or 'private' schools, we mean the recognised private schools only. For the purposes of this paper, we exclude PA schools and compare the case of pure 'government schools' with the case recognised private unaided schools which we call simply 'private schools' or 'PUA schools'.

Table 1 compares the quality of PUA and government schools at the district level over the decade 1992-2002, using some commonly used quality measures. In general, PUA schools tend to have better infrastructure (pucca building, access to drinking water and toilets) than government schools; however, government schools have significantly narrowed the gap in this respect over the decade. Compared to recognised PUA schools, Government schools have significantly higher pupil-teacher ratio (more than double at the primary level) and the situation does not change much over the decade under consideration. Furthermore, recognised PUA schools employ a higher proportion of female teachers. Although Table 1 does not show this, compared to government schools, PUA schools also have younger teachers, fewer teachers with (pre-service) training and fewer vacant teaching positions (see Pal 2010). Thus despite significant public

[9] However, in most states, there are no board examinations at the primary of upper primary levels so there is no strong incentive for private schools to seek government 'recognition' except if the school wishes ultimately to become a secondary school and affiliate with an exam board.

[10] Even the District Information System on Education (DISE) data collection exercise – introduced in the late 1990s – does not have the mandate/authority to collect information on all unrecognised PUA schools. Thus, even today there is no way of reliably knowing the number of unrecognised PUA schools in India, though see Kingdon (2008) for various estimates.

interventions over the 1990s (for instance, the District Primary Education Programme) to improve government schools, input differences between recognised PUA and government schools persisted by 2002.

2.1. Growth of private schools

Using 6th and 7th AISES data, we first calculate the average share of recognised PUA schools in total schools at a given level (primary, upper primary and secondary) in a district, and also the corresponding district literacy rates, as summarised in Table 2. In each of the two years, the share of private schools at the secondary level (e.g., 15% in 1992) is significantly higher than at the primary level (e.g., 4.4% in 1992) [11]. Over the course of the decade 1992-2002, the pace of private school growth gathered momentum at all levels, with private school share at secondary level reaching 28% in 2002.

We also examine the nature of private school growth at primary, upper primary and secondary levels *across the regions* in our sample. This is shown in Table 3. As shown in Table 2, the share of recognised PUA schools is significantly higher at the secondary level (relative to primary and upper primary levels) over the period 1992-02. Table 3 highlights pronounced inter-regional variation in the rate of private school growth. We classify all districts into five regions, namely, east (Assam, Bihar, Orissa, West Bengal (WB)), west (Gujarat and Maharashtra), north 1 (Punjab and Haryana), north 2 (Madhya Pradesh (MP), Rajasthan and Uttar Pradesh(UP)) and south (Andhra Pradesh (AP), Karnataka, Kerala and Tamil Nadu (TN)). In general, the rate of private school growth is relatively lower in the eastern states, especially at the primary and the upper primary levels. At the secondary level, the highest share of PUA schools is found in the socially backward northern states (in the North 2 region) namely MP, Rajasthan and UP, which are generally known for failing government schools (see Dreze and Kingdon, 2001).

[11] It should be borne in mind that at the primary school level the share of private schools in total schools appears lower than the true private share because there are no data on the private unrecognised schools. Kingdon (2008) shows estimates suggesting that a high proportion of private schools at the primary level remain unrecognised. Thus, our estimate of the private share of total primary schools is an underestimate.

2.2. Literacy rates and gender gap

Unfortunately AISES data do not provide information on any learning outcomes. Hence we combine 1992 and 2002 AISES data with age/gender specific literacy data available from the 1991 and 2001 Census data respectively. Our analysis focuses on children aged 10-19 years. This choice has been guided by the fact that we could not obtain literacy rates for primary school age children 5-9 years old. While 10-14 literacy rates correspond broadly to literacy rates for upper primary level of education, those for 15-19 correspond to that for the secondary level. We also analyse the rate of growth of literacy rate for 10-19 years old taken together, and we do so for both male and female children. As before, we classify our sample into five regions, namely, east, west, north1, north 2 and south; this allows us to consider private school share and literacy rates not only for the whole of India, but also for the sub-regions in our sample.

Table 2 shows the literacy progress at primary, upper primary and secondary levels between 1992 and 2002, while Table 3 presents the male and female literacy rates for 10-14 and 15-19 age groups across the regions. Not surprisingly, literacy rates are lower for female children, in both the 10-14 and 15-19 age groups. The gender difference is significantly higher in the worse performing regions, e.g., see eastern (comprising of Assam, Bihar, WB and Orissa) and northern zones 2 (comprising of UP, MP and Rajasthan). Compared to the national average, age/gender specific literacy rates are lower in these two regions and higher in the west, south and north 1 (Punjab and Haryana) regions.

3. EMPIRICAL ANALYSIS

As set out in the introduction, we empirically model literacy rates as well as the gender gap in literacy rates, among children aged 10-19 years old. We also split 10-19 years old children into upper-primary (10-14 years) and secondary (15-19 years) school age groups and repeat the analysis separately for each age group. This allows us to explore the difference, if any, in the estimates between upper-primary and secondary school age group children.

3.1. Modelling universal literacy and gender gap

Our central objective pertains to the determination of literacy rates (L_{ilt}) and gender gap in literacy rates (G_{ilt}) in the *i*-th district for a particular schooling level *l* in year *t*; l refers to upper primary (10-14), secondary (15-19) or the pooled category (10-19 years) while t refers to 1992 and 2002. We start with the most general specification as follows:

$$L_{ilt} = \alpha P_{ilt} + \quad \gamma' X_{ilt} + \nu_{1il} + \mu_{1lt} + \varepsilon_{1ilt} \qquad (1)$$

$$G_{ilt} = \beta P_{ilt} + \delta' X_{ilt} + \nu_{2il} + \mu_{2lt} + \varepsilon_{2ilt} \qquad (2)$$

Here P_{ilt} is the share of private unaided schools (in total schools) at the *l*-th school level in district *i* at time *t*. In addition we control for other possible factors affecting literacy and gender gap in literacy rates in our sample. In particular, the set of control variables *X* includes adult literacy rates, share of urban population, proportion of scheduled caste and scheduled tribe population, ratio of female to male child 0-6 year olds and also supply of schools per child at the given level with a view to minimise the omitted variable bias as far as possible. Given the close link between literacy and earnings, we consider adult literacy rates to be a good proxy for income or wealth. Since scheduled caste and scheduled tribe population are more disadvantaged than the general population and are also over-represented in Indian poverty, these SC and ST variables would also proxy for poverty. Since urban literacy rates are often much higher than the rural literacy rates in the Indian context, we include share of urban to rural population as a proxy for urbanisation with a view to explore its effects on literacy and gender gap. Son preference may also play an important role in parental allocation of resources for education and other accounts. In the absence of a better alternative, our measure of son-preference is the district ratio of surviving female to male children in the 0-6 age-range. Furthermore, it could be that in districts where there are more private schools, the overall supply of schooling is greater and that, access to schooling is greater. If so, the private school effect could capture an effect of schooling availability. In order to eliminate this possibility, we include number of total schools (at the relevant level, upper-primary, secondary or both pooled) per 100 children as an additional explanatory variable.

Given the multi-level nature of our data, we include both district (ν_1, ν_2) and year-specific (μ_1, μ_2) fixed effects respectively in equations (1) and (2); the remaining errors are captured by (ε_1, ε_2), which are independently and identically distributed. Use of panel data fixed effects models allow us to obtain consistent estimates net of unobserved heterogeneity. In particular, both private school growth and literacy may be influenced by some unobserved factors like district's culture, institutions, labour market characteristics, gender and caste relations; the resultant estimates would be biased if these unobserved factors are correlated with the error term, thus justifying the use of panel data fixed effects models. Given that we have only two data points for each district, these fixed effects estimates are also equivalent to the underlying first difference estimates of changes in literacy in terms of changes in private school share and also changes in other X variables (that eliminates time-invariant unobserved heterogeneity).

A further option is to determine changes in literacy and gender gap in literacy over the period 1992-2002 at a given level (upper primary, secondary or both pooled) as functions of 1992 levels of private school share and other X variables as follows:

$$\Delta L_{il} = \varphi + \beta P_{il} + \delta' X_{il} + \xi_{1il} \qquad (3)$$

$$\Delta G_{il} = \varphi + \beta P_{il} + \delta' X_{il} + \xi_{1il} \qquad (4)$$

Thus, in order for reverse causality to bias our estimates, private investors need to anticipate future literacy ten years in advance, which could be ruled out without much loss of generality. Since we have data for only two years, after differencing we get one time period for each cross section unit (i.e., district) and hence we omit the time subscript. In this case however we cannot convincingly redress the potential bias, if any, arising from unobserved heterogeneity.[12]

Perhaps a better alternative for us is to make use of fixed effects instrumental variable (FE-IV) method that can redress the bias generated by both unobserved heterogeneity and simultaneity, which to date remains the most widely accepted

[12] Note that, we do not estimate difference in difference estimates in this case, as they are identical to the FE level estimates of literacy and gender gap when we have only two-period panel data.

approach.[13] The crucial thing here is to identify the relevant instrumental variables, which are likely to affect private school share, but would be directly uncorrelated with the unobserved error terms in equations (1) and (2). Identification of a convincing instrumental variable could be contentious and we rely on existing literature to sort it out. The existing literature suggests that private school growth is induced by the poor quality of local government schools (Muralidharan and Kremer, 2006), and also by the availability of or access to local public infrastructure including access to road, rail network, post office, telegraph office (Pal 2010). While quality of local government schools is likely to be a direct determinant of district literacy rates, districts' access to local public infrastructure should not affect literacy rates directly. While it is possible to quibble with its validity, we use the district's local public infrastructure as an instrument for the private school share in total district enrolment. Thus we instrument the private school share variable P in equations (1) and (2) by the district's access to public transport and communications infrastructure, which makes the equation exactly identified. We expect the direction of simultaneity bias to be negative since districts with higher literacy rate tend to have lower share of private schools (Muralidharan and Kremer, 2006; Pal, 2010). In an attempt to minimise the bias arising from the unobserved heterogeneity, we also include both district and year specific fixed effects in each equation. We compare FE-OLS estimates with FE-IV estimates with a view to identify the nature of the estimation bias.

Finally, we compare the full sample estimates for each level of schooling (upper-primary and secondary) with estimates obtained from possible sub-samples. To this end, we re-estimate equations (1) and (2) for various sub-samples: (a) districts located in northern and southern states; (b) districts targeted by the District Primary Education Programme DPEP; (c) SC/ST children (10-14 and 15-19 years old).

In each case, we cluster the standard errors at the district level, which corrects for any correlation in errors within a state; otherwise OLS estimates would be biased.

[13]Alternatively, economists have also attempted to make use of experiments that objectively randomize treatments to asses their effects in the presence of unobserved heterogeneity. These randomized experiments too are subject to criticisms that they lack generalizability and often do not adhere to the requirement of treatment randomness.

4. RESULTS

4.1. Full sample estimates

4.1.1. FE OLS estimates

We start in Table 4 with simple fixed effects linear models for literacy and gender gap equations (1) and (2).[14] All standard errors are robust to clustering at the district level.

It follows that higher private school share is associated with significantly higher literacy for all age groups while it is associated with significantly lower gender gap in literacy only among 10-14 year old children. Cleraly, the literacy effect of private school growth is most pronounced for the younger age-group, 10-14 year olds, who naturally benefitted more from the recent trend of private school growth around the country. In addition, adult literacy rates (25-49 years) tend to boost literacy and lower gender gap in literacy in all the relevant age groups (10-19 years old) that we consider. It also follows that districts with higher share of ST population experienced higher literacy during 1992-2002 while the effect of SC population has generally been insignificant in our sample. Rate of urbanisation however fails to have any significant effect on literacy.

Table 5 augments the estimates shown in Table 4 by including two characteristics of existing government schools which are likely to affect literacy. In general, exclusion of these two possible explanatory variables are likely to generate an underestimation bias in the effect of private school share, as districts with better state schools are likely to have lower share of private schools. This under-estimation bias is confirmed in the private school effect shown in Table 5. Compared to Table 4, the size of the private school effect of literacy as well as gender gap in literacy is greater in Table 5 and also it holds for all the age groups under consideration.

It has often been argued that the necessity of being accountable to parents causes private schools and teachers to apply more effort. The notion that private management of schools leads to higher teacher effort is supported in some recent study on India. For instance, using data from 20 Indian states Muralidharan and Kremer (2008) find that within the same village, teacher absence rate in private schools is about 8 percentage points lower than in government schools. This is similar to the findings in Kingdon and Banerji

[14] Corresponding pooled OLS estimates are shown in Appendix Table A2.

(2009) for Uttar Pradesh and Bihar. More generally, our findings of a limited positive private school effect on cognitive skills are consistent with a growing body of literature that finds similarly, using data from different sources and using different methods (Desai, et al, 2008; French and Kingdon, 2010).

4.1.2. Estimates of changes in literacy

While these FE estimates minimises any potential bias arising from unobserved heterogeneity, they tend to ignore the problem of simultaneity bias. One option is to consider the changes in literacy and gender gap in literacy equations (3) and (4). So we next consider robust ols estimates of models (3) and (4) in terms of lagged values of all explanatory variables, as summarised in Table 6. Much in line with the FE OLS level estimates, these first difference estimates suggest a favourable effect of private school growth on literacy for all age groups of our interest. In addition, these estimates tend to highlight a favourable effect of private school growth on gender gap in that higher private school share is associated with lower gender gap over the decade 1992-2002 for all age groups 10-14, 15-19 and 10-19 year olds. Note also that compared to the level effects of private school growth as shown in Tables 4 and 5, these first order effects are somewhat smaller. While one can argue that the potential simultaneity arising from private school growth is minimised here as we use one period (i.e., a decade) lagged value of private school share while determining changes in literacy and gender gap, questions may still arise about the potential bias of these estimates arising from unobserved heterogeneity that we try to address in terms of FE-IV estimates discussed below.

4.1.3. FE IV estimates

In order to reduce the biases arising from both unobserved heterogeneity and simultaneity, we next estimate the fixed effects instrumental variable (FE-IV) estimates. Following Pal (2010), we consider the district's access to public transport and communications infrastructure as measured by percentages of villages in the district with access to these local public infrastructural goods (available from Census data) as our instrument for the district's private school share. While this variable is closely correlated to the district

private school share, it is unlikely to directly affect literacy or gender gap in literacy. Appendix Table A3 shows the relationship between district's access to private schools (at various levels) and that to public transport and communication infrastructure (with and without other controls). There is evidence that districts with better access to public transport and communication infrastructure tend to have higher access to private schools at upper-primary, secondary levels and also both levels pooled together.

Accordingly, we use district's access to public infrastructure to as an instrument to determine literacy. The resulting equation is exactly identified so that we do not need to report the Sargan-Hansen statistic of over-identification. These FE-IV estimates are summarised in Table 7. While the literacy effect is significant for 10-14 year olds, it fails to be so for 15-19 year olds; for 10-19 year olds the effect is nearly significant at 10% level. The size of the private school effects on both literacy and gender gap in literacy is larger in Table 7 than in Table 5 and it holds for both the age groups concerned. FE-IV results show that correcting for this simulataneity bias raises the coefficients on private school share. In other words, FE-OLS model was under-estimating the private school effect of literacy. In general, these estimates highlight a substantial impact of private school growth on literacy, though its effect on gender gap in literacy remains rather weak as before. This conforms our expectation, as the simultaneity bias is likely to be negative here: districts with higher literacy are likely to have lower share of private schools (Muralidharan and Kremer, 2008; Pal, 2010).[15] As before, compared to 15-19 year olds, literacy effect is stronger for the younger age-group 10-14 year olds, who are likely benefit more from recent growth of private schools around the country.

While FE-IV fixed effects estimates could provide a more stringent approach than ordinary FE OLS, identification of the right instruments could be questioned. Our FE-IV estimates tend to suggest that FE-OLS estimates tend to suffer from an under-estimation bias though the direction of the relationships remain unchanged. Moreover, given that there are only two years in our panel, these FE-OLS estimates are identical with the first difference estimates, which eliminate unobserved heterogeneity and also minimises

[15] Sample correlation between the district literacy and private school share are -0.19, -0.02 and -0.11 respectively at upper primary (10-14), secondary (15-19) and pooled 10-19 level.

reverse causality. Accordingly we consider the FE-OLS estimates as a lower bound of the true estimates of private school effect on literacy and gender gap in literacy in our sample - the rest of our analysis in this paper focuses on FE-OLS estimates only.

4.2. Sub-sample estimates

4.2.1. A comparison between the northern and the southern districts

Given the pronounced heterogeneity among the Indian states, one may wonder whether private school effect of literacy and gender gap may vary across the Indian states. To test this proposition, we compare fixed effects OLS estimates of (1) and (2) among 10-19 year olds in the four north Indian states, namely, Bihar, MP, Rajasthan and UP with those for the southern states (AP, Karnataka, Kerala and Tamil Nadu). Table 8 suggests that, even after controlling for son preference and all other factors, the private school effect of literacy and gender gap do vary between these two regions: while private school share remains insignificant to determine both literacy and gender gap among 10-19 year olds in the northern districts, both these effects are significant in the southern districts. Compared to the southern districts, districts located in the large north Indian states are not only poorer on average, but also tend to have worse gender relations so that private school growth in the northern regions tends to have a rather limited impact. One possibility is that those who choose private schools in the northern districts are likely to be relatively better-off and more motivated to do well at schools and would have become literate even in the absence of private school growth in the region. Ceteris paribus, private school growth fails to have a perceptible impact on universal literacy in the northern (as opposed to southern) region.

4.2.2. Case of DPEP districts

It may be possible to argue that our results are confounded by the fact that the District Primary Education Programme (DPEP) ran in India from mid 1990s onwards. We know that educationally backward districts received the DPEP programme that focused on boosting literacy in the district. If these are also the districts with higher private school shares in total schools, then there may be a positive correlation between DPEP and private

school share, across districts. If the DPEP raised literacy rates then the 'private school share' variable would 'pick it up' and thus generate an upward bias on the private share variable. Alternatively, if DPEP districts – being poorer – had smaller private school shares, then DPEP and private share would be negatively correlated and any positive private school effect would be biased downwards. The correlation between DPEP and private school share varies across the levels of schooling: while the correlation is only about 0.0480 at the primary level, it is around 0.19 for both upper primary and secondary levels.

Schmid (2006) analysed the impact of the DPEP programme in the Indian districts, focusing mainly on the school participation (schooling access, rather than literacy) outcome. Using data from Schmid (2006), we have created a binary variable 'DPEP', which takes a value of 1 if DPEP had been adopted in a district any year between1994-2000 and a value of zero otherwise. Accordingly, we select the districts that implemented DPEP in any year (between 1994 and 2000) to re-estimate the FE OLS models of literacy (1) and gender gap in literacy (2).[16] These results are summarized in Table 9. It is evident that private school share is significant for literacy among all three age groups though, as before, the effect is most pronounced among younger 10-14 year old children. If we compare these estimates with those in Table 5 (for all children), it is evident that the private school effect of literacy is less pronounced for children in the DPEP districts and it holds for all the age groups considered. This is a further confirmation that private school effect of universal literacy is more pronounced in better-off regions. As before, private school effect of gender gap in literacy remains insignificant for all age groups under consideration.

4.2.3. Universal literacy among SC/ST children

One may also argue that the effect of private school growth on overall literacy may blur what happens to the marginalised group of low caste population, who are far behind the general population in terms of literacy achievements. We use 2001 Census literacy data

[16] We consider DPEP districts rather than using DPEP as a binary variable in the full sample as this variable is not time-varying for the two sample years 1991 and 2001 (it was implemented in between these years) and we prefer the FE estimates to other specifications as discussed earlier.

for SC/ST children aged 10-19 years old to examine if there is any private school effect for these marginalised social groups. Since separate SC/ST literacy data is not available for 1991 and also since 2001 SC/ST literacy data does not offer separate information for male and female children in any age group, we can only obtain OLS estimates (with robust standard errors, correcting for clustered errors) of literacy rates for this sub-sample as summarized in Table 10. What is evident from these estimates is that the private school effect of literacy is significant for both 10-14 and 15-19 years old SC children; for ST children however the private school effect is significant only for younger 10-14 year olds, as before. Given that these are simple ols estimates, we compare these marginal private school effects on literacy with those for all children (Appendix Table A2); it follows that marginal private school effect for SC and ST children are significantly higher than the general population and also it holds for both 10-14 and 15-19 year olds; true effects are likely to be even larger when we address the likely underestimation bias arising from both unobserved heterogeneity and simultaneity. In other words, there is suggestion that there are some large literacy gains to be had from private school growth even among SC/ST children, especially among 10-14 year olds. As before, there is no suggestion that private school growth can influence SC/ST gender gap in literacy in our sample.

5. CONCLUDING COMMENTS

Since early 1990s, there has been a rapid growth of private schooling in India. Despite some recent attempts to compare the efficiency of public and private schools in certain regions in India in recent years, the implications of private school growth for literacy rates and gender gap in literacy rates, two key educational MDGs, remain rather unexplored. Using two period district-level panel data (1992-2002), compiled from various official sources, the present paper attempts to bridge this gap in the literature.

In the light of the available data, our analysis focuses on children aged 10-19 years old, i.e., those likely to attend upper primary and secondary schools. Among various possible alternatives, we compare FE-OLS estimates with first difference models and also FE-IV estimates with a view to minimise the bias arising from both simultaneity and

endogeneity in our sample. There is some evidence that the true effects of private school shares on literacy and gender gap are likely to be underestimated if one ignores these potential biases.

Results highlight that growing share of private schools exerts a pronounced effect on literacy for children in our sample, but its effect on gender gap in literacy remains rather limited, if at all. In general, districts with greater share of private schools tend to experience significantly higher literacy, which is in line with the household/child-level evidence that private schools are more efficient in imparting learning. Further estimates highlight interesting regional variation between northern and southern regions in India: while the literacy effect of private school growth is statistically significant in the southern region, it is not so in the districts located in the large north-Indian states, namely, Bihar, MP, Rajasthan and UP. Also compared to all sample districts, positive literacy effect of private school growth is much weaker in DPEP districts. There are also some intra-group variations (within a district) in the literacy effect of private school growth: compared to the general population, the effect is significantly larger for 10-14 year old SC/ST children in our sample.

One drawback has been that the paper used information on recognised private schools only while recent evidence suggests a faster growth of unrecognised private schools (at the primary and junior school levels) often catering to children from relatively less well off households (Desai et al. 2008). There is yet to be a study of relative efficiency of recognised and unrecognised private (vis-à-vis state) schools. It would therefore be interesting to explore whether the results of this paper hold, when one takes account of both recognised and unrecognised private schools in India.

References

Bashir, S. (1994) 'Public versus Private in Primary Education: Comparison of School Effectiveness and Costs in Tamil Nadu', PhD Dissertation, London School of Economics.

Becker GS (1981) 'Altruism in the family and selfishness in the market place', *Economica* 48(189):1–15.

Bedi, A.S. and A. Garg. (2000) 'The Effectiveness of Private versus Public Schools: The Case of Indonesia', *Journal of Development Economics* 61 pp. 463-94.

Beegle, K and D. Newhouse. (2006). The Effect of School Type on Academic Achievement: Evidence from Indonesia', *Journal of Human Resources* XLI(3):529-557.

Bloom, J.E., D.E. Cohen and M.B. Malin. 2006 'Educating All Children: A Global Agenda', Cambridge MA: MIT Press.

Brown, D. (1979) 'The Location Decision of the Firm: An Overview of Theory and Evidence' *Papers in Regional Science*, 43(1), 23-39.

Census of India. (2001) 'Data Highlights: Migration Tables D1, D2 and D3', Office of the Registrar General, India, Government of India. www.censusindia.net/results/dseries/data_highlights_D1D2D3.pdf.

Desai, S., A. Dube, R. Vanneman and R. Banerji. (2008) 'Private Schooling in India – A New Landscape', mimeo NCAER.

Drèze, J and H. Gazdar. (1997) 'Uttar Pradesh: The Burden of Inertia', in Jean Dreze and Amartya Sen (edited) 'Indian Development: Selected Regional Perspectives', New Delhi: OUP.

Drèze, J. and G.a Kingdon. (2001) "School Participation in Rural India", *Review of Development Economics*, 5, No. 1: 1-33.

Duflo, Esther and Rema Hanna (2005) "Monitoring Works: Getting Teachers to Come to School", NBER working paper 11880.

Epple, D. and R. Romano. (1998) 'Competition between Private and Public Schools, Vouchers and Peer Group Effects', *American Economic Review* 88(1) pp. 585-610.

French, Rob and Geeta Kingdon (2010) "The relative effectiveness of private and government schools in Rural India: Evidence from ASER data", mimeo, Institute of Education, University of London.

Glewwe, P. (2002) 'Schools and Skills in Developing Countries: Education Policies and Socio-economic Outcomes', *Journal of Economic Literature*, 40(2), pp.436-482, 2002.

Hanushek, E. and L. Woessman. (2008) 'The Role of Cognitive Skills in Economic Development' *Journal of Economics Literature* 46(3), 607-668.

Hansen, N. (1965). 'Unbalanced Growth and Regional Development', *Western Economic Journal* 4, pp. 3-14.

Holmlund, H. and K. Sund. (2008) 'Is the gender gap in school performance affected by the sex of the teacher?'Labour Economics 15(1), p. 37-53.

Hoxby, C. (1994) 'Do Private Schools Provide Competition for Public Schools?' NBER working paper No. 4978d.

Kingdon, Geeta. (1996) "The Quality and Efficiency of Public and Private Schools: A Case Study of Urban India", *Oxford Bulletin of Economics and Statistics*, 58, No.1: 55-80.

Kingdon. G. (2008) "School-Sector Effects on Student Achievement in India", Chapter 6 in Chakrabarti, R. and P. Peterson (eds.). *School Choice International: Exploring Public-Private Partnerships*, MIT Press

Kingdon, G. and R. Banerji (2009) "Addressing School Quality: Some Policy Pointers from Rural North India", RECOUP Policy Briefing Number 5, Faculty of Education, University of Cambridge.

Kingdon, Geeta & Teal, Francis, 2010. "Teacher unions, teacher pay and student performance in India: A pupil fixed effects approach," *Journal of Development Economics*, Elsevier, vol. 91(2), pages 278-288,

Muralidharan, K. and M. Kremer. (2006) 'Public and Private Schools in Rural India', Chapter 5 in Chakrabarti, R. and P. Peterson (eds.). *School Choice International: Exploring Public-Private Partnerships*, MIT Press.

Pal, S. (2004). 'How Much of the Gender Difference in Child School Enrolment Can be Explained?

Evidence from Rural India', *Bulletin of Economic Research* 56(2), pp. 133-158.
Pal, S. (2010). 'Public Infrastructure, Location of Private Schools and Primary School Attainment in an Emerging Economy' *Economics of Education Review*, online first.
PROBE Team. (1999) '*Public Report on Basic Education in India*', Oxford University Press.
Psacharopoulos, George (1988) "Education and Development: A Review," *World Bank Research Observer*, Oxford University Press, vol. 3(1), pages 99-116, January.
Reinikka, R. and J. Svensson. (2002) 'Coping with Poor Public Capital', *Journal of Development Economics*, 69, pp. 51-69.
Tooley, J. (2004) 'Private Education and Education for All', Economic Affairs, December, Blackwell publishing.
Watkins, K. 2004. Private Education and Education for All – Or How Not to Construct an Evidence Based Argument: A Reply to Tooley', Economic Affairs, December, Blackwell publishing.

Table 1. A comparison of government and private unaided schools, 1992-2002

	Private schools		Govt. schools	
	1992 Mean (SD)	2002 Mean (SD)	1992 Mean (SD)	2002 Mean (SD)
% of female teachers in total				
Primary	0.55 (0.26)	0.55 (0.24)	0.35 (0.23)	0.38 (0.15)
Upper primary	0.50 (0.27)	0.50 (0.24)	0.30 (0.21)	0.32 (0.17)
Secondary	0.45 (0.26)	0.44 (0.22)	0.28 (0.20)	0.24 (0.19)
% of low caste teachers				
Primary	0.09 (0.13)	0.11 (0.13)	0.22 (0.19)	0.24 (0.19)
Upper primary	0.08 (0.13)	0.10 (0.12)	0.17 (0.14)	0.23 (0.17)
Secondary	0.07 (0.11)	0.08 (0.08	0.15 (0.13)	0.17 (0.12)
% of schools with pucca building	0.78 (0.17)	0.77 (0.39)	0.66 (0.24)	0.79 (0.34)
% of schools with lavatory	0.66 (0.23)	0.71 (0.22)	0.33(0.26)	0.41 (0.27)
% of schools with drinking water	0.84 (0.17)	0.91 (0.13)	0.58 (0.24)	0.78 (0.17)
Pupils per teacher				
Primary	30.7 (12.5)	34.3 (31.6)	39.1 (16.2)	67.1 (70.5)
Upper Primary	30.8 (11.2)	20.6 (50.0)	31.5 (11.5)	35.3 (58.3)
Secondary	29.1 (10.9)	13.7 (23.0)	28.1 (8.4)	29.7 (19.1)

Source: 6^{th} and 7^{th} AISES data

Note: PUA schools refer to private unaided schools. Government schools do not include private aided schools.

Table 2. Pace of private school growth and of youth literacy in the Indian districts

	Average share of recognised PUA schools in all schools Mean (sd)		**10-19 literacy rates**			
	1992-93	**2002-03**	**1992-93**		**2002-03**	
			Male	Female	Male	Female
Primary	0.044	0.08	---	---	---	---
	(0.07)	(0.10)				
Upper primary	0.11	0.17	0.77	0.58	0.87	0.77
	(0.14)	(0.18)	(0.13)	(0.21)	(0.11)	(0.17)
Secondary	0.15	0.28	0.74	0.51	0.84	0.69
	(0.15)	(0.22)	(0.12)	(0.21)	(0.11)	(0.18)

Source : AISES and Census Data.

Table 3. Regional variation in literacy and private school share: Means and standard deviations for the (1992 and 2002) pooled data

Level	East	West	North-west	North	South	All
			Mean share (sd) of recognised private schools (in total schools)			
Primary	0.003(0.007)	0.084 (0.11)	0.05 (0.07)	0.12 (0.09)	0.047 (0.07)	0.07 (0.09)
Upper primary	0.023 (0.04)	0.058 (0.08)	0.16 (0.17)	0.28 (0.15)	0.096 (0.11)	0.15 (0.16)
Secondary	0.10 (0.12)	0.22 (0.12)	0.18 (0.15)	0.35 (0.23)	0.22 (0.15)	0.22 (0.20)
			Mean literacy (sd) rates			
Female 10-14	0.53 (0.17)	0.82 (0.12)	0.81 (0.11)	0.55 (0.20)	0.80 (0.17)	0.66 (0.21)
Male 10-14	0.71 (0.13)	0.910.05)	0.89 (0.06)	0.78 (0.13)	0.89 (0.09)	0.81 (0.13)
Female 15-19	0.47 (0.17)	0.74 (0.14)	0.73 (0.14)	0.46 (0.19)	0.72 (0.20)	0.59 (0.21)
Male 15-19	0.70 (0.12)	0.88 (0.07)	0.85 (0.08)	0.76 (0.12)	0.84 (0.12)	0.79 (0.13)
Female 10-19	0.54(0.17)	0.78 (0.13)	0.77 (0.13)	0.51 (0.19)	0.76 (0.18)	0.63 (0.21)
Male 10-19	0.72 (0.12)	0.89 (0.06)	0.87 (0.07)	0.77 (0.13)	0.86 (0.11)	0.80 (0.13)

Source: 6th and 7th AISES data and 1991 and 2001 Census data.

Note: Indian regions: south=AP, Tamil Nadu, Kerala, Karnataka; West: Gujarat, Maharashtra; East: Assam, Bihar, Orissa, West Bengal; North-west=Punjab, Haryana; North=Madhya Pradesh, Rajasthan, Uttar Pradesh.

Table 4. Fixed effects OLS: Determinants of the district literacy rate and of the gender gap in literacy rate

	Literacy			Gender gap		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	10-19	10-14	15-19	10-19	10-14	15-19
Private school Share	0.146***	0.255***	0.0718**	-0.0668	-0.136***	-0.0166
	(0.0351)	(0.0427)	(0.0319)	(0.0426)	(0.0373)	(0.0432)
Literacy rate 25-49	0.848***	0.837***	0.817***	-0.595***	-0.537***	-0.626***
	(0.0553)	(0.0621)	(0.0562)	(0.0539)	(0.0493)	(0.0702)
Urbanisation	-0.0172	-0.0671	0.0465	-0.0218	-0.00171	-0.0536
	(0.0945)	(0.0742)	(0.121)	(0.0703)	(0.0498)	(0.0982)
SC	0.0837	0.0741	0.120	-0.0829	-0.115	-0.0707
	(0.0833)	(0.0718)	(0.135)	(0.160)	(0.0895)	(0.249)
ST	0.159*	0.246***	0.0703	0.0984	0.0179	0.183*
	(0.0846)	(0.0821)	(0.0941)	(0.0827)	(0.0759)	(0.0980)
Female-male ratio 0-6	0.455**	0.576***	0.323	0.0846	0.0670	0.104
	(0.214)	(0.217)	(0.215)	(0.226)	(0.204)	(0.279)
Schools per 100 children	-1.564	1.236	-9.426	0.600	-1.260	6.831
	(8.323)	(5.088)	(15.12)	(5.090)	(3.195)	(9.153)
Constant	-0.161	-0.255	-0.0438	0.386*	0.375**	0.387
	(0.199)	(0.201)	(0.202)	(0.209)	(0.189)	(0.263)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	606	606	606	607	607	607
R-squared	0.860	0.853	0.849	0.707	0.771	0.575
Number of districts	353	353	353	353	353	353

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5. FE-OLS estimates of the extended model:
Determinants of the district literacy rate and of the gender gap in literacy rate

	Literacy			Gender gap		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	10-19	10-14	15-19	10-19	10-14	15-19
Private school Share	0.245***	0.369***	0.141***	-0.0765*	-0.0938*	-0.0517
	(0.0587)	(0.0649)	(0.0465)	(0.0422)	(0.0562)	(0.0362)
Gov sch: drinking water	0.0567**	0.0523**	0.0581**	-0.0173	-0.0181	-0.0162
	(0.0247)	(0.0246)	(0.0270)	(0.0208)	(0.0212)	(0.0261)
Gov sch: pucca building	0.107***	0.141***	0.0675***	0.00697	-0.00312	0.0186
	(0.0225)	(0.0249)	(0.0232)	(0.0238)	(0.0289)	(0.0253)
Literacy rate 25-49	0.801***	0.773***	0.797***	-0.515***	-0.461***	-0.562***
	(0.0627)	(0.0718)	(0.0619)	(0.0527)	(0.0557)	(0.0612)
Urbanisation	0.106	0.0116	0.220**	-0.106*	-0.0590	-0.158*
	(0.0865)	(0.0799)	(0.0947)	(0.0641)	(0.0477)	(0.0958)
SC	0.148*	0.155**	0.163	-0.0908	-0.112	-0.0808
	(0.0820)	(0.0723)	(0.132)	(0.161)	(0.0948)	(0.248)
ST	0.0828	0.169*	0.00836	0.0935	0.00453	0.174
	(0.0801)	(0.0964)	(0.0729)	(0.0990)	(0.0903)	(0.118)
Female-male ratio 0-6	0.398	0.449*	0.314	-0.297	-0.157	-0.438
	(0.252)	(0.258)	(0.256)	(0.245)	(0.195)	(0.396)
Schools per 100 children	-0.0426	-0.0013	-0.1313	0.0316	0.0003	0.1062
	(0.0819)	(0.0505)	(0.1517)	(0.0459)	(0.0303)	(0.0835)
Constant	-0.244	-0.280	-0.159	0.728***	0.566***	0.888**
	(0.231)	(0.238)	(0.233)	(0.231)	(0.181)	(0.377)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	513	513	513	514	514	514
R-squared	0.858	0.849	0.851	0.724	0.752	0.604
Number of discode	317	317	317	317	317	317

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6. OLS estimates of first difference model: Determinants of literacy rate and of the gender gap in literacy rate

	Change in literacy			Change in gender gap		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	10-19	10-14	15-19	10-19	10-14	15-19
Private school share	0.106***	0.144***	0.0632**	-0.332***	-0.300***	-0.256***
	(0.0348)	(0.0347)	(0.0300)	(0.0480)	(0.0371)	(0.0480)
Gov sch: drinking water	0.0355	0.0397*	0.0132	-0.107***	-0.0649***	-0.137***
	(0.0231)	(0.0238)	(0.0225)	(0.0253)	(0.0200)	(0.0323)
Gov sch: pucca building	0.0138	0.0343	0.00938	0.00897	-0.0291	0.0403
	(0.0261)	(0.0302)	(0.0249)	(0.0334)	(0.0301)	(0.0387)
Literacy rate 25-49	-0.295***	-0.365***	-0.225***	0.142***	0.186***	0.0931*
	(0.0409)	(0.0434)	(0.0433)	(0.0374)	(0.0333)	(0.0474)
urbanisation	-0.0350	0.000670	-0.0627	0.211***	0.146***	0.238***
	(0.0346)	(0.0325)	(0.0384)	(0.0617)	(0.0464)	(0.0741)
SC	0.229***	0.288***	0.167***	-0.0849	-0.130**	-0.0914
	(0.0637)	(0.0722)	(0.0558)	(0.0589)	(0.0546)	(0.0707)
ST	0.0435	0.0474	0.0427	0.0991***	0.106***	0.0786**
	(0.0303)	(0.0351)	(0.0284)	(0.0300)	(0.0258)	(0.0378)
Female-male ratio 0-6	0.105	0.319**	-0.106	-0.124	-0.241**	0.0653
	(0.122)	(0.125)	(0.126)	(0.127)	(0.112)	(0.156)
Schools per 100 children	0.0599	0.03000	0.1300	0.0104	-0.05098*	0.1912
	(0.0832)	(0.0505)	(0.1273)	(0.0585)	(0.0283)	(0.1311)
Constant	0.0578	-0.150	0.260**	0.0289	0.145	-0.157
	(0.126)	(0.128)	(0.128)	(0.128)	(0.112)	(0.158)
Observations	291	291	291	298	298	298
R-squared	0.385	0.454	0.290	0.426	0.515	0.307

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 7. FE-IV Determinants of district literacy rate and of the gender gap in literacy rate

	Literacy			Gender gap		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	10-19	10-14	15-19	10-19	10-14	15-19
Private school share	0.299*	0.453**	0.184	0.214	0.239	0.219
	(0.179)	(0.228)	(0.173)	(0.194)	(0.206)	(0.211)
Gov sch: drinking water	0.0554**	0.0529**	0.0541**	0.00376	-0.00296	0.0106
	(0.0239)	(0.0238)	(0.0254)	(0.0236)	(0.0215)	(0.0313)
Gov sch: pucca building	0.106***	0.137***	0.0697**	-0.00564	-0.0216	0.00965
	(0.0286)	(0.0320)	(0.0282)	(0.0282)	(0.0291)	(0.0346)
Literacy rate 25-49	0.817***	0.782***	0.817***	-0.474***	-0.469***	-0.480***
	(0.0674)	(0.0658)	(0.0806)	(0.0663)	(0.0597)	(0.0986)
urbanisation	0.00981	-0.0581	0.106	-0.104	-0.0288	-0.189*
	(0.0758)	(0.0751)	(0.0876)	(0.0739)	(0.0675)	(0.106)
SC	0.119	0.141	0.110	-0.100	-0.108	-0.100
	(0.134)	(0.144)	(0.138)	(0.132)	(0.131)	(0.169)
ST	0.139	0.212*	0.0535	0.0721	-0.0187	0.170
	(0.101)	(0.109)	(0.102)	(0.0994)	(0.0987)	(0.125)
Female-male ratio 0-6	0.438*	0.438	0.374	-0.467*	-0.380	-0.565*
	(0.252)	(0.290)	(0.246)	(0.247)	(0.261)	(0.298)
Schools per 100 children	0.0720**	0.0603**	0.1016*	-0.0402	-0.0451*	-0.0095
	(0.0362)	(0.0283)	(0.0608)	(0.0356)	(0.0256)	(0.0745)
Constant	-0.291	-0.277	-0.224	0.850***	0.764***	0.944***
	(0.229)	(0.267)	(0.227)	(0.224)	(0.241)	(0.275)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	510	510	510	511	511	511
Number of districts	316	316	316	316	316	316

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8. FE OLS: Determinants of district literacy rate and of the gender gap in literacy rate for northern and southern regions

	(1)	(2)	(3)	(4)
	North		South	
	Bihar, MP, Rajasthan, UP		AP, Karnataka, Kerala, Tamil Nadu	
VARIABLES	Literacy 10-19	Gender gap 10-19	Literacy 10-19	Gender gap 10-19
Private school share	0.0150	0.0579	0.236*	-0.154*
	(0.0650)	(0.0878)	(0.138)	(0.0789)
Gov sch: drinking water	0.0700**	0.100**	0.0202	-0.0204
	(0.0331)	(0.0419)	(0.0377)	(0.0308)
Gov sch: pucca building	0.115***	-0.0156	0.109**	0.0119
	(0.0345)	(0.0471)	(0.0424)	(0.0473)
Literacy rate 25-49	0.899***	-0.675***	1.074***	-0.474***
	(0.0554)	(0.0623)	(0.133)	(0.0999)
urbanisation	-0.0997	0.344	0.0102	-0.145***
	(0.139)	(0.364)	(0.0816)	(0.0406)
SC	-0.160**	-0.0869	0.184	-0.267***
	(0.0655)	(0.265)	(0.152)	(0.0880)
ST	0.336***	0.512	0.399	0.169
	(0.0965)	(0.427)	(0.387)	(0.249)
Female-male ratio 0-6	0.198	-0.605	1.703***	-0.238
	(0.273)	(0.548)	(0.500)	(0.342)
Schools per 100 children	11.15**	4.373	6.253	-2.829
	(4.367)	(6.992)	(6.339)	(3.675)
Constant	-0.0571	0.923*	-1.631***	0.699**
	(0.261)	(0.496)	(0.450)	(0.312)
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	182	183	146	146
R-squared	0.971	0.847	0.872	0.716
Number of districts	120	120	77	77

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 9. FE OLS : Determinants of district literacy rate and of the gender gap in district literacy rate (DPEP districts)

	Literacy			Gender gap		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	10-19	10-14	15-19	10-19	10-14	15-19
Private school share	0.218***	0.316***	0.120*	-0.0288	-0.0934	0.0175
	(0.0742)	(0.0796)	(0.0626)	(0.0643)	(0.0709)	(0.0587)
Gov sch: drinking water	0.0454	0.0280	0.0554	-0.0112	-8.10e-05	-0.0197
	(0.0391)	(0.0376)	(0.0434)	(0.0320)	(0.0335)	(0.0371)
Gov sch: pucca building	0.0903***	0.123***	0.0567*	0.0176	0.0176	0.0199
	(0.0308)	(0.0314)	(0.0337)	(0.0344)	(0.0438)	(0.0410)
Literacy rate 25-49	0.821***	0.828***	0.798***	-0.596***	-0.582***	-0.600***
	(0.0922)	(0.111)	(0.0841)	(0.0791)	(0.0906)	(0.0847)
Urbanisation	-0.257*	-0.287*	-0.194	0.236	0.0389	0.406
	(0.146)	(0.161)	(0.149)	(0.213)	(0.123)	(0.397)
SC	-0.0104	-0.00266	-0.0375	0.0954	0.0308	0.163
	(0.122)	(0.119)	(0.139)	(0.130)	(0.112)	(0.172)
ST	0.258*	0.398***	0.130	0.103	0.000972	0.208
	(0.143)	(0.152)	(0.153)	(0.141)	(0.149)	(0.147)
Female-male ratio 0-6	0.584	0.737*	0.372	-0.618	-0.196	-0.999
	(0.395)	(0.417)	(0.387)	(0.407)	(0.309)	(0.718)
Schools per 100 children	0.0832**	0.0756**	0.0686	-0.0592	-0.0457*	-0.0666
	(0.0336)	(0.0336)	(0.0512)	(0.0370)	(0.0241)	(0.0943)
Constant	-0.368	-0.526	-0.137	0.971**	0.593**	1.302*
	(0.389)	(0.407)	(0.377)	(0.382)	(0.297)	(0.668)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	242	242	242	243	243	243
R-squared	0.916	0.911	0.909	0.790	0.823	0.652
Number of districts	146	146	146	146	146	146

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 10. 2002 OLS: Determinants of district literacy rate in the SC/ST population

	(1)	(2)	(3)	(4)
	SC literacy		ST literacy	
VARIABLES	10-14 yrs	15-19 years	10-14 yrs	15-19 years
Private school share	0.308***	0.226***	0.623***	0.295
	(0.0541)	(0.0506)	(0.196)	(0.195)
Gov sch: drinking water	-0.106*	-0.0811	0.0296	0.229
	(0.0576)	(0.0550)	(0.186)	(0.179)
Gov sch: pucca building	0.0199	-0.0660	0.122	-0.176
	(0.0576)	(0.0540)	(0.153)	(0.154)
Literacy rate 25-49	0.981***	1.055***	0.0874	-0.225
	(0.0706)	(0.0570)	(0.215)	(0.195)
Urbanisation	0.121***	0.0928*	-0.0753	0.0291
	(0.0444)	(0.0474)	(0.174)	(0.185)
SC	-0.137	-0.420***	0.686**	1.006***
	(0.117)	(0.135)	(0.326)	(0.311)
ST	0.308***	0.296***	0.353*	0.202
	(0.0546)	(0.0561)	(0.192)	(0.184)
Female-male ratio 0-6	-0.519**	-0.538***	0.801	1.045
	(0.209)	(0.184)	(0.760)	(0.747)
Schools per 100 children	0.0557	-0.0086	-0.4417**	0.0815
	(0.0338)	(0.0792)	(0.1701)	(0.3172)
Constant	0.691***	0.701***	-0.487	-0.645
	(0.213)	(0.190)	(0.749)	(0.715)
Observations	216	216	216	216
R-squared	0.617	0.710	0.150	0.091

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix

Table A1. Descriptive Statistics

Label	Obs	Mean	Std. Dev.	Source
Private school share, 6-9	828	0.0626884	0.087236	AISES
Private school share, 10-14	828	0.1382488	0.1605384	AISES
Private school share, 15-19	828	0.2152241	0.2013952	AISES
Private school share, 10-19	828	0.9311594	0.2533359	AISES
Enrolment, 10-14	791	0.3985718	0.2708459	AISES
Enrolment, 15-19	791	0.2711397	0.2245323	AISES
Enrolment, 10-19	791	0.3348558	0.1998333	AISES
Literacy rate 10-14	760	0.744034	0.1604009	Census
Literacy rate 15-19	760	0.6961569	0.1574186	Census
Literacy rate 10-19	760	0.7200955	0.1579778	Census
Literacy gap 10-14	828	0.1278763	0.111892	Census
Literacy gap 15-19	828	0.170923	0.1340042	Census
Literacy gap 10-19	828	0.1493997	0.1215899	Census
Literacy 25-49 years	828			
Govt. school with pucca building	766	0.8444108	0.1648442	AISES
Govt. school with drinking water	763	0.6807499	0.2352175	AISES
Female teachers in govt. UP school	795	0.3125357	0.1918513	AISES
Female teachers in govt. sec. school	819	0.2966186	0.1898494	AISES
Teaching vacancies in gov sch. 10-14	795	0.1130191	0.2421194	AISES
Teaching vacancies in govt. 15-19	793	0.3120095	0.479444	AISES
Poverty rate	769	0.2271017	0.1220107	NSS
Share of SC population	760	0.1647903	0.0874591	Census
Share of ST population	730	0.0929837	0.1530782	Census
% of villages with public transport and communications	784	24.49912	17.55916	Census
Southern region	828	0.1908213	0.393186	AISES
Western region	828	0.1183575	0.3232262	AISES
Eastern region	828	0.2512077	0.4339699	AISES
Northern region 1	828	0.0724638	0.259411	AISES
Northern region 2	828	0.3671498	0.4823192	AISES

Table A2. Pooled OLS: Determinants of district literacy rate and gender gap in literacy rate

	Literacy rates			Gender gap in literacy		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	10-19	10-14	15-19	10-19	10-14	15-19
Private school share	0.150***	0.131***	0.100***	-0.0209	-0.00905	-0.0179
	(0.0242)	(0.0271)	(0.0211)	(0.0326)	(0.0312)	(0.0275)
Gov sch: drinking water	-0.00804	-0.0428**	0.0179	-0.00158	0.0186	-0.0221
	(0.0167)	(0.0177)	(0.0152)	(0.0164)	(0.0155)	(0.0179)
Gov sch: pucca building	-0.0536***	-0.0475**	-0.0559***	0.117***	0.114***	0.123***
	(0.0187)	(0.0213)	(0.0170)	(0.0228)	(0.0217)	(0.0246)
Literacy rate 25-49	0.965***	0.926***	1.000***	-0.501***	-0.473***	-0.535***
	(0.0275)	(0.0316)	(0.0253)	(0.0262)	(0.0249)	(0.0298)
Urbanisation	-0.0664***	-0.0403*	-0.0703***	-0.0329	-0.0397*	-0.0314
	(0.0244)	(0.0242)	(0.0260)	(0.0246)	(0.0204)	(0.0292)
SC	0.133***	0.212***	0.0842**	-0.0268	-0.0664	-0.000806
	(0.0385)	(0.0450)	(0.0343)	(0.0468)	(0.0430)	(0.0537)
ST	0.0566**	0.0793***	0.0323	-0.0487*	-0.0513**	-0.0573*
	(0.0243)	(0.0261)	(0.0267)	(0.0278)	(0.0255)	(0.0311)
Female-male ratio 0-6	-0.347***	-0.321**	-0.365***	-0.331**	-0.280*	-0.360**
	(0.131)	(0.153)	(0.123)	(0.144)	(0.145)	(0.153)
Schools per 100 children	-0.00103	0.0297	-0.0442	0.0389	0.0155	0.1509***
	(0.0512)	(0.0285)	(0.0949)	(0.0288)	(0.0167)	(0.0495)
year02	-0.0672***	-0.0473***	-0.0793***	0.0246***	0.00574	0.0450***
	(0.00785)	(0.00863)	(0.00787)	(0.00801)	(0.00708)	(0.00986)
east	0.00694	-0.00610	0.00773	-0.0656***	-0.0490***	-0.0875***
	(0.00929)	(0.0105)	(0.0113)	(0.0128)	(0.0124)	(0.0137)
west	0.0821***	0.0993***	0.0579***	-0.0868***	-0.0747***	-0.103***
	(0.0127)	(0.0129)	(0.0128)	(0.0130)	(0.0132)	(0.0133)
south	0.0552***	0.0745***	0.0296***	-0.0714***	-0.0563***	-0.0915***
	(0.00992)	(0.0115)	(0.0111)	(0.0111)	(0.0114)	(0.0113)
north1	0.0776***	0.106***	0.0464***	-0.131***	-0.116***	-0.155***
	(0.0114)	(0.0134)	(0.0130)	(0.0180)	(0.0155)	(0.0222)
Constant	0.582***	0.586***	0.578***	0.692***	0.601***	0.764***
	(0.125)	(0.145)	(0.116)	(0.135)	(0.138)	(0.142)

R-squared	0.914	0.896	0.918	0.758	0.759	0.728

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Observations = 514 districts

Table A3. Justification of using district's access to public infrastructure as an instrument

	Share of private schools in total schools					
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	10-19	10-14	15-19	10-19	10-14	15-19
Access to infrastructure	0.00530***	0.00338***	0.00722***	0.00454***	0.00298***	0.00610***
	(0.000494)	(0.000405)	(0.000688)	(0.000628)	(0.000503)	(0.000904)
Monthly Per Capita Expenditure				0.000357***	0.000158**	0.000556***
				(0.000101)	(6.85e-05)	(0.000164)
SC				0.355**	0.0824	0.627**
				(0.178)	(0.163)	(0.253)
ST				-0.238	-0.135	-0.341
				(0.153)	(0.101)	(0.232)
urban				4.34e-08*	2.20e-08	6.47e-08*
				(2.24e-08)	(1.61e-08)	(3.61e-08)
Constant	0.0505***	0.0578***	0.0432**	-0.117***	0.00116	-0.236***
	(0.0121)	(0.00991)	(0.0169)	(0.0452)	(0.0354)	(0.0679)
Observations	784	784	784	652	652	652
R-squared	0.281	0.196	0.252	0.341	0.213	0.318
Number of discode	407	407	407	355	355	355

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1