Leading education and social research
Institute of Education
University of London

Do "Child Friendly" Practices affect Learning? Evidence from Rural India

Sushmita Nalini Das

Department of Quantitative Social Science

Working Paper No. 14-03
February 2014

Disclaimer

Any opinions expressed here are those of the author(s) and not those of the Institute of Education. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

DoQSS Workings Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Department of Quantitative Social Science, Institute of Education, University of London
20 Bedford Way, London WC1H 0AL, UK

Do "Child Friendly" Practices Affect Learning? Evidence from Rural India

Sushmita Nalini Das[1]

Abstract

This paper examines the effects of "child-friendly" practices on learning in rural India. These are a set of pedagogical practices intended to improve education outcomes by increasing children's inclusion in their learning environment. They are widely promoted in international development circles, and are an increasingly important plank of Indian education policy. This paper offers the first quantitative evidence of their impact. Data is drawn from a survey of 12,576 primary school pupils in government schools in rural India. Incidence levels of six pedagogical practices, each representing a different aspect of child-friendliness described in Indian policy documents, are drawn from high-quality classroom observations. Estimates of their impact on low-stakes reading and maths test-scores are then generated using a school fixed effects value-added model. The main finding is that child-friendly practices, while well-intentioned, generally have insignificant effects on test-scores. Even in circumstances where the practices show some effects, they do not always have the positive impact which would be expected based on their popularity in policy discourse. Further, no strong evidence emerges that the practices differentially affect the test-scores of high and low ability pupil groups. These findings highlight substantial flaws in the content of the publically available evidence base which underlies important elements of Indian education policy, and the gains which could be made from more rigorous analysis at the policy formulation stage.

JEL classification: I21, I28

Keywords: Child-Friendly practices, National Curriculum Framework, test-scores, primary education, value-added models, India

[1] Department of Quantitative Social Science, Institute of Education, University of London (sdas@ioe.ac.uk)

Acknowledgements:

Thank you to the ESRC, UK for funding my PhD. research of which this paper is a part of. Special thanks to my PhD. supervisors Prof. Lorraine Dearden and Dr. John Jerrim for valuable guidance and comments throughout the research. Additional thanks to Dr. Paul Atherton, Prof. Emla Fitzsimons and Dr. Shenila Rawal for helpful comments. Finally, additional thanks also to Dr. Suman Bhattacharjea, Dr. Wilima Wadhwa and Dr. Rukmini Banerji at the ASER Centre, in New Delhi, for their encouragement and granting me access to the datasets. Any errors are my own. Comments welcome.

1. Introduction

This paper examines the effects of "child-friendly" teaching practices on primary school pupils' test-scores in rural India. Child-friendly practices are a set of teaching and pedagogical practices intended to improve education outcomes by increasing children's inclusion in their learning environment. They have become an increasingly important plank of education policy in recent years, particularly in India, where they are actively encouraged by policymakers. However, no formal evaluation of their impact on pupil's test scores has yet been carried out. This paper offers an initial exploration of their impact, using data on 12,576 primary school pupils in rural India.

The concept of child-friendly environments has been used extensively by the United Nations Children's Fund (UNICEF), in the context of rehabilitating conflict zones, and has extended to the education setting more generally. They are now actively recommended by education policy makers in a number of developing countries, including India. In the last decade, increasing effort has been expended to promote these practices in primary education. UNICEF's *Child Friendly Schools Manual* (2009) lays out steps to set up child-friendly environments within schools. The manual stresses the importance of "quality" of teaching and learning, and of practices that are focused on "…the child as learner [who] is central to the process of teaching and learning". The sections on "Organizing classrooms and learning spaces", "Pedagogic Materials" and "Learning and Teaching Methods" (*Pg. 16 to 26*) stress the importance of having classrooms that are "stimulating", with pedagogic materials that are "appropriate teaching and learning aids", and teachers who use methods that are "child-centred, interactive methodologies [that] make learning enjoyable and exciting to students….." It is expected that such child-friendly environments will improve pupil retention and academic performance. The manual cites extensively from experiences of teachers, pupils and parents in laying out and justifying its recommendations.

This concept of child-friendly education appears throughout recent Indian education policy documents. The *National Curriculum Framework 2005* (NCF) lays out a framework for central and

state governments to develop their curriculums, text-books and standards for teaching practices, as recommended by the government's education policy advisory organization *National Centre for Education Research and Training* (NCERT). The NCF explicitly recommends the use of practices along child-friendly lines. Chapter 4 of the framework, "*School and Classroom Environment*", highlights positive and negative teaching behaviours and practices, includes descriptions of a suitable "physical environment" within schools, and suggestions for the development of "nurturing and enabling environments" within the classroom. It also stresses the importance of the "participation and [respectful inclusion] of all children" and gives advice on the appropriate use of "learning resources" for teaching *(Pg.78)*.

The 2009 Right-to-Education Act, which is the most important piece of Indian education policy of recent years, draws extensively from the NCF. One prominent reform it makes is the introduction of a mandatory eligibility test for teachers, alongside minimum education qualifications. To qualify, sections of the test require teachers to display "practical" knowledge of the education process, including the appropriate use of child-friendly practices contained in the NCF[2]. This test is now administered by all Indian state education authorities.

Despite the increasing importance being assigned to these practices in Indian education policy, no formal evaluation of their effects has yet been published[3]. The documents cited make very specific recommendations to teachers about their methods and practices, without providing any quantitative evidence on the effectiveness of such practices. It is important to examine whether the faith placed in them is reasonable, that is, whether they have a measurable impact on pupil test-scores.

There are a limited amount of papers within the economics of education literature that directly examine the effectiveness of teaching practices on learning. The motivation for doing so has been to re-direct focus away from standard teacher characteristics (such as qualifications, training and

[2] See sample questions within preparatory materials for the teaching eligibility test: http://www.teachers-eligibility-test.com/online-practice/child-development-and-pedagogy-online-mock-test/

[3] The state of Karnataka notes that there has been an evaluation of the Joyful Learning program which advises child-friendly practices, but this is not publically available.

experience) towards teaching practices in the classroom. Prominent papers include: Aslam and Kingdon (2011) who use self-reports from teachers on time spent using various teaching practices; Lavy (2011) who use pupil reports on their teachers practices; and Kane et al (2010) who use longitudinal data on teaching practices collected via classroom observations in a sample of American schools.

Separately, the recent surge of activity in randomised evaluations has included papers which examine the impact of interventions which alter school resources and/or teacher incentives on measures of teaching practices. Papers include: Glewwe et al (2003) & Muralidharan and Sundaraman (2009) who examine whether performance bonuses incentivise teachers to use more effective teaching methods; Duflo et al (2009) who examine whether teachers hired on temporary contracts in Kenyan schools taught more "actively" than colleagues on permanent contracts; and Muralidharan and Sundaraman (2010a) who study the effect of providing report cards to teachers in India that highlight learning weaknesses of pupils, accompanied with suggestions of effective teaching practices.

None of these papers use direct measures of child-friendly practices as either input or outcome variables. This is the first paper to introduce measures of these practices into the education production function and examine their effect on pupil test-scores. The paper builds on the work done within the "*Child-Friendly Classrooms*" section in Chapter 4 of the report *"Inside Primary Schools: Teaching & Learning in Rural India."* (Bhattacharjea et al, 2011)[4]. This chapter was the first to identify markers of child-friendly practices, and examined their impact on test-scores using descriptive statistics. This paper uses the same six markers of child friendly practices as the Inside Primary Schools report.

The analysis uses a lagged value-added specification of the education production function. It accounts for prior-attainment, school, teacher and family characteristics. Data on the incidence of

[4] See Pg. 42 of publication Bhattacharjea S., W. Wadhwa, R. Banerji, "*Inside Primary Schools: Teaching and Learning in Rural India*", Pratham Mumbai Education Initiative, 2011. I was a part of the research team and contributed to this publication. Pratham's ASER centre have kindly supplied the data used in this paper.

child-friendly practices is drawn from high-quality independent classroom observations in a sample of rural Indian primary schools spread over 15 districts across the states of Andhra Pradesh, Rajasthan, Assam, Jharkhand and Himachal Pradesh. Results from a more causal analysis method, show that child-friendly practices, as articulated in Indian policy documents, have little effect on pupil test scores. This finding underscores the importance of using evidence rigorously when setting educational policy. It suggests that a policy which has been introduced with limited evaluation may well be ineffective, and that Indian policymakers may need to supplement or further refine child-friendly strategies in order to maximise educational performance. The conclusion argues that improved support mechanisms (eg. reducing teacher absenteeism, improving the stability of teaching responsibility environments within schools) may be needed in order for child-friendly practices to become significant drivers of pupil test scores.

The paper is organized as follows; section 2 discusses the data, section 3 discusses the background to the value-added framework and sets up the empirical strategy, section 4 discusses the results and section 5 concludes.

2. Data

2.1 The "Inside Primary Schools" Surveys

Data for this paper are drawn from the Inside Primary School (IPS) surveys. The surveys aimed to create a longitudinal dataset capturing the test-scores and educational circumstances of primary school pupils from different administrative, social and geographical contexts within India. Data collection was not associated with any intervention at the surveyed schools. The surveys were fielded between 2009 and 2010, by the ASER Centre. One state was purposively selected from each major geographical region in India: Northeast (Assam), North (Himachal Pradesh), Central

(Jharkhand), West (Rajasthan) and South (Andhra Pradesh). Three rural districts were then purposively selected within each state[5].

Individual pupils within the districts were selected into the study population by a two-stage sampling process. At the first stage, 60 government schools with primary grades were sampled in each district from the most recent list of government schools for the district, using probability proportional to size (PPS) weighted by total enrolment in Grades 1 to 5. In the second stage, up to 50 pupils were randomly sampled from each school: up to 25 each from the enrolment registers of grades 2 and 4. Second stage sampling was done by surveyors on arrival at the school. Many selected schools had fewer than 25 pupils enrolled in the targeted grades. In these cases, all pupils enrolled in the concerned grade were included. The final datasets contain an average of 12 pupils in each of grades 2 and 4 per school (i.e. average of 24 observations per school)[6].

Three survey rounds were conducted. The surveys collected information on pupils' learning levels in reading and maths, school infrastructure, teacher characteristics, teaching & learning practices and households. Table A.1 in the appendix presents key information collected within each of these rounds. Baseline and final test-scores, are available for 22,000 pupils across grades 2 and 4.

[5] The districts were selected by representatives from the Ministry of Human Resource Development, UNICEF and ASER Centre. A number of factors such as economic and education characteristics along with feasibility of financial & logistical arrangements within each of these districts were taken into consideration. The sampling techniques and study location places limits on the external validity of the results. For example, the findings in this paper apply to government school children in rural India. Private schools in rural India and urban locations might yield different results. For example, government schools in urban locations can reasonably be expected to have better management, training resources and involvement by the local community and education authorities. Therefore it might be that these factors increase the incidence of child-friendly practices in urban government-schools and they might very well be significant drivers of learning in this context. With this in mind, one should refrain from generalizing this paper's results beyond rural government schools.

[6] There is a concern that exists in the literature (Kane & Staiger (2002)) that small samples within schools increases the risk of high-variability in test-scores within schools, increasing the incidence of measurement error which results in a "noisy" error term and subsequently imprecise coefficient estimates. Strategies are outlined in the literature to correct for this sampling variation (see Aaronson et al, 2007). In this paper, the child-friendly practice coefficient estimates are not statistically significant to begin with. The correction is going to increase the standard errors within the models and would be especially important to do if the estimations generated positive findings bordering on statistical significance, which is not the case here.

2.2 Selection of Child-Friendly Practices

At Round 1, thirty minute classroom observations of teaching practices were conducted. The incidence of a range of teaching and learning practices were recorded. Of these, six practices explicitly discussed in the NCF and representing child-friendly pedagogies have been selected for analysis.

1. **Displaying children's work in the classroom**

   The NCF recommends that "Classrooms can be made lively by displaying children's work on the classroom walls as well as in different parts of the school…" with the expectation that this will "….attract and retain children…." to school *(Section 4.1, Paragraph 2, Pg. 79).*

2. **Teachers smile, laugh or joke with students**

   Laughing, smiling or joking is included to capture a friendly and enjoyable environment. The NCF states that "Children will learn only in an atmosphere where they feel they are valued... The association of learning with fear, discipline and stress, rather than enjoyment and satisfaction, is detrimental to learning."*(Section 4.2, Paragraph 1, Pg. 82)*.

3. **Pupils encouraged to ask teachers questions during lessons**

4. **Teachers use local information to contextualize text-book content**

   The third and fourth practices examine participatory approaches to learning. The NCF criticizes the frequent use of teaching practices that enforce discipline or classroom management, and encourages children to freely engage with their teachers and ask questions: "Rules regarding maintaining silence in classrooms, answering 'one at a time' and answering only if you know the right answer can undermine the values of equality and equal opportunity. Such rules may also discourage practices that are integral to children's learning." (*Section 4.4, Paragraph 2, Pg. 87*). The NCF also favours teaching practices that encourage "The participation of the community in the child's world of education and learning…" (*Section 4.5, Paragraph 2, Pg. 88*). It suggests that a good approach to incorporate this is by "Influenc[ing] the content of subjects [text-book

material] and add[ing] local, practical and appropriate examples" (*Section 4.5, Paragraph 2, Pg. 88*).

5. **Teacher use teaching & learning materials (TLM) beyond text-books during lessons**
6. **Teachers organize small group activities during lessons**

The fifth and sixth practices examine the use of learning resources other than text-books. The NCF posits that "Teaching aids and other materials, as well as books, toys and games help make school interesting for children". (*Section 4.6.7, Paragraph 1, Pg 94*). The NCF stresses that exclusive funding to schools has been made available for the purchase of a variety of teaching aids, and encourages teachers to explore their use. The NCF also encourages the shift from "vertically organized classrooms which assume that all children are being addressed by the teacher together and that they are all at the same stage and are all expected to do the same thing". (*Section 4.6.1, Paragraph 5, Pg 90*), and instead recommends the use of small-group activity.

The six practices were recorded as binary variables, with '1' indicating that a practice was observed and '0' indicating that it had not been used. All visits to schools were unannounced, and teachers were not aware in advance that the survey was to be conducted or of its contents. This survey structure limits the probability that practices observed were not part of normal operation within the classrooms visited.

Table 1 presents the incidence of the six practices. 20% of classrooms were observed as displaying at least one of the child-friendly practices. The most frequently observed practice is "pupils asking teachers questions" and the least frequent practice observed is "teachers used teaching & learning materials other than the textbook", across both Grade 2 and 4 classrooms. Variables indicating the "frequency" or "intensity" with which practices were observed are not available. Due to the large number of classrooms covered within each district, it was cost-effective to use observation schedules that only allowed for whether a practice was observed or not. The impact of the binary structure of the practice variables on results is discussed in Section 3.3.1 and 3.3.2.

Table 1: Classrooms with Child-Friendly Teaching Practices

	GRADE 2 Room	GRADE 4 Room
Classrooms where pupils work was displayed	22%	28%
Classrooms where teachers smiled/laughed/joked with pupils	21%	24%
Classrooms where teachers used localised info. to contextualize lessons	23%	21%
Classrooms where pupils encouraged to ask teachers questions	26%	32%
Classrooms where teachers used TLMs other than the textbook	15%	12%
Classrooms where teachers divided pupils into small group activities	16%	17%

2.3 Data Set-up

Classroom observations of teaching practices have been matched to test-scores, pupil characteristics and teacher & school characteristics across survey rounds. Two important decisions have been taken:

(1) Matching teachers to pupils in the presence of flexible teaching responsibilities

All lower primary teachers that were present on the day of survey during Round 1 were interviewed. 2501 such teacher interviews were completed. Of these, 1364 teachers taught grade 2 and/or grade 4. Within this set, 366 teachers gave overlapping reports of their teaching responsibilities i.e. more than one teacher in a school reported teaching grade 2 and/or grade 4. For example, in one case 3 teachers reported teaching the same group of grade 2 pupils: one teacher taught them reading & maths; a second taught them only reading; a third taught only maths.

"Flexible teaching responsibilities", in which a single grade (and even a single subject) is taught by multiple teachers over an academic year, is entirely plausible in the Indian context. Previous research has repeatedly shown that rural Indian schools face chronic teacher absenteeism, rising as high as 25% of the work-force within schools. (ASER surveys 2006-2011, Kremer et al, 2005). This is

corroborated in the IPS data, in which an average of 25% of teachers were absent across all 3 rounds of data collection (Bhattacharjea et al, 2011). In such a situation, it is not surprising to see instability of teaching responsibility, with teachers filling in for each other frequently. There is very limited literature available on modelling the influence of flexible teaching responsibilities on pupil achievement. Strategies that can be drawn from the literature (though they have not been applied for this purpose) include averaging teachers' characteristics (Kingdon 2006) or restricting analysis only to those schools with only one teacher reporting teaching a particular grade (Metzler & Woessman, 2012).

This paper conducts analysis on two data set-ups. These set-ups are generated from teaching responsibility reports from both the teacher self-reports and head-teacher. Teachers were asked to self-report their teaching responsibilities at round 1, while at round 3 head-teachers were asked to report which of their teacher's most frequently taught each grade over the course of the academic year. The first set up includes schools in which a single teacher is identified as holding "primary" or "sole" responsibility for teaching a particular grade (i.e. pooling "flexible" and "stable" teaching responsibility schools). The "flexible" schools here include cases where both teacher and head-teacher reports match, but other teachers in the school have also reported some responsibility for teaching a particular grade. It is essentially the "most frequent" teacher of the grade's characteristics that has been retained for analysis, but it is important to remain aware that other teachers in the school have influenced final test score. Therefore, as a robustness check, the a second set up is chosen to include only those schools where a single teacher can be identified as holding the sole responsibility for teaching a particular grade and exact matches of teachers to pupils can be made (i.e. only "stable" teaching responsibility schools). This set up strictly includes cases where both head-teacher and teacher reports match, no other teachers in the school report teaching the grade, and no teachers were absent from the school (and thus unable to report their teaching responsibilities) during Round 1. The impact of this is discussed in more detail in the results section.

(2) Use of teaching practices to explain performances in both reading and maths

Due to time constraints, classroom observations were made for either reading or maths during data-collection. In cases where schools had the same teacher for grade 2 (or 4) reading and maths, it is assumed that the teacher applied these practices for both subjects. This seems reasonable given that the child-friendly practices selected for analysis are behaviour-based and are not subject-specific. However, the practices will not vary across the two subjects and a "pupil fixed-effects" specification cannot be applied. In cases where schools had a separate teacher for reading and maths, only test-scores in the subject for which classroom observations were observed are retained.

The final sample retained for the analysis includes 6564 and 6192 pupils in grades 2 and 4 respectively, across 645 schools taught by 858 teachers. Within this, 3138 and 2883 pupils in grade 2 and 4 respectively, across 375 schools taught by 496 teachers form the "stable" school sub-sample. Approximately 60% of the pupils in both grades attended classrooms in which at least one Child-friendly practice was observed. Sample description statistics in appendix table A.2 show that the characteristics of the sub-sample(s) used in this paper are similar to the main sample of 22,000 pupils in the IPS surveys.

2.4 Learning Outcomes

Test-scores are used to measure learning outcomes. In India there are no mandatory national or state achievement exams during primary school. Test-scores in the IPS surveys have been drawn from reading and maths tests which were independently designed by the survey administrators, the ASER Centre. The same tests were administered at baseline (Round 1) and end line (Round 3), approximately a year apart. The tests were based directly on text-book material in use within each of the state's primary schools. Efforts were made to keep tests consistent with a state's individual learning goals and also comparable across states. Approximately, half of the material in the tests was drawn from the current grade's text-books, while the other half was drawn from material from

the 2 previous grades. Including material from previous grades in the tests allows for a more accurate analysis of each pupil's level of knowledge upon entering (and exiting) the grade under observation.

Table 2 presents the mean, standard-deviation and range of the test-scores attained across the survey rounds. Striking points that emerge from the table are; first, compared to the range of the tests, the spread of scores is quite wide in both subjects and for both grades, as evidenced from the standard deviations; second, most pupils do not attain full-marks on the final tests.

Table 2: Test-Scores in the IPS data

	Grade 2	**Grade 4**
Baseline Reading (raw scores)	Mean = 13.1 S.D. = 9.6 Range: 0 to 34	Mean = 3.6 S.D. = 3.0 Range: 0 to 12
Final Reading (raw scores)	Mean = 17.5 S.D. = 10.1 Range: 0 to 34	Mean = 4.8 S.D. = 3.1 Range: 0 to 12
Baseline Maths (raw scores)	Mean = 6.7 S.D. = 3.5 Range: 0 to 13	Mean = 9.4 S.D. = 5.3 Range: 0 to 21
Final Maths (raw scores)	Mean = 7.9 S.D. = 3.4 Range: 0 to 13	Mean = 10.7 S.D. = 5.2 Range: 0 to 21

Note: Scores presented are only for the pupils used in the paper (i.e. 6564 and 6192 grade 2 and 4 pupils respectively). For the full IPS sample, refer Table A.2 in the appendix.

Although the tests were pitched at levels that state curriculums required pupils to attain by Grade 2 and 4, many pupils obtained round 3 scores that were far below the maximum. These tests were "low-stake"[7] for the pupils and the main tasks in the reading tests included reading simple two-letter words, sentences & paragraphs and writing dictated alphabets & two letter words. Tasks in the maths tests included number recognition, simple 2 digit computations and word-problems. These tasks mark key learning goals in the lower primary grades (according to the states' text-

[7] "Low-stake" test results are typically not associated with any publication or public dissemination of individual pupil/school performance. They are also not usually linked to any rewards or sanctions for the school management or teachers on the basis of the test-results (Muralidharan and Sundaraman, 2010a).

books)[8]. Within the grade 2 samples, only 42% pupils could read simple two-letter words and only 72% could solve 1 digit addition problems at round 3 - tasks expected of them at the end of grade 1. Within the grade 4 samples, only 31% of pupils could read a paragraph from a text taken from grade 2 text-books at round 3, and only 85% could solve 2 digit addition problems intended for grade 2. These statistics match the main IPS survey samples. In general, the surveys found pupils with learning achievements of at least 2 grades behind the expectation.

These findings are consistent with the literature on learning levels of Indian primary school pupils. The nationally representative ASER surveys have routinely found severe under-achievement among rural primary school pupils. In 2010, only 35% of grade 4 pupils could read text intended for grade 2 and only 23% could solve division problems intended for grade 3 across the country. Even lower achievements are seen in the 2007-2008 SchoolTELLS surveys [9]. Only 20% of grade 4 pupils sampled in Uttar Pradesh could read text intended for grade 2 and solve two digit addition problems. For a representative sample of schools in Andhra Pradesh, the APREStt surveys find that less than half of grade 2 pupils unable to do single digit addition learn it by grade 5. Among other reasons, these low achievements are often attributed to 'curriculum mismatch' (i.e. over-ambitious curriculums designed without accounting for the practicalities of time constraints, teacher-skill levels, and varying pupil learning trajectories) and lack of test-evaluations within primary school grades (Pritchett & Beatty, 2012 and Banerji, 2000).

2.5 Test-Scores and Child-Friendly Practices

Table 3 presents raw test-scores for both grades by whether or not at least one child friendly practice was observed during classroom observations[10]. For both the grade 2 and grade 4 samples, on average pupils obtained higher test-scores in classrooms that had been observed as having at least one child-friendly practice in survey rounds 1 and 3. This difference in scores is highly

[8] See Chapter 1 of Bhattacharjea et al (2011).
[9] The SchoolTELLS surveys are the precursor to the IPS surveys in the Indian states of Bihar and Uttar Pradesh. Very similar tests and survey methods were used for these surveys.
[10] Table A.3 in the appendix shows descriptions of variables used for analysis. Data is categorized in a similar way as in Table 3.

significant. The effect of having a child-friendly practice seems to have had a smaller impact on round 3 test-scores with a slightly higher impact on reading than for maths for grade 2. In general, for both grades, the raw data suggests that child-friendly practices have a positive impact on pupil learning.

Table 3: T-tests of difference in test-scores by whether child-friendly practices observed

	GRADE 2			GRADE 4		
	No Practice	With at least one practice	Diff.	No Practice	With at least one practice	Diff.
Baseline Reading (raw scores)	11.9	13.7	-1.8***	3.1	4	-0.9***
Final Reading (raw scores)	16.1	18.3	-2.2***	4.3	5.1	-0.8***
Baseline Maths (raw scores)	6.2	7	-0.8***	8.4	10	-1.6***
Final Maths (raw scores)	7.7	8	-0.3*	10.1	11.2	-1.1***

Note: ***p < 0.01 **p < 0.05 *p<0.01; these statistics are generated for the full-sample (i.e. pooling the "flexible" and "stable" schools).. For the t-tests, standard errors are clustered at the school level

In further analysis z-scores have been created separately for subject and for each grade, to render the achievements across both subjects and grades comparable. The overall mean score (across schools) in a particular subject for a particular grade is subtracted from a pupil's actual score in that subject and divided by the standard deviation in that subject for the relevant grade. Thus each pupil's performance on the tests is relative to the average in each subject across schools, but within his/her grade. Figure 1 in the appendix, shows the pooled kernel density distributions of standardized test-scores across grade 2 and 4.

3. Method

Table 3 indicated that exposure to child-friendly practices are associated with higher test-scores. However, this does not mean that the relationship is necessarily causal. This section develops an empirical model which attempts to identify the causal impact of child-friendly practices.

3.1 The Value-Added Model

The main approach that is taken to estimate the impact of education inputs on pupil test-scores within the economics of education literature is the specification of a production function. Ideally, to estimate the impact of various inputs on a pupil's test score, we should have a complete history of all inputs from birth up to the point in time that the test-score has been generated. Since such data is generally not available, researchers have developed empirical representations of the education process that require lesser data inputs, but also account for biases that occur due to incomplete information.

A common approach to account for unobserved past inputs to the education process is the use of a lagged test-score. The inclusion of a lagged test-score can account for historical characteristics not captured by contemporaneous measurements of family and school inputs[11]. It also is usually thought of as capturing innate abilities (which are typically not observable) of pupils to an extent. Both of these features help mitigate biases stemming from incomplete information on the education process. An econometric specification that contains lagged test-scores is referred to as a "value-added" model, as it examines the "value" added by the inputs in the period between the lagged and contemporaneous test score. Evidence on input effects generated by papers which include a lagged test-score generally argue themselves to be more reliable than papers which do not include a lagged test-score (Hanushek et al (2003 and 2005), Clotfelter et al (2007), Aaronson et al (2007) and Slater et al (2012)).

[11] "Contemporaneous" refers to the set of inputs used in the period between the baseline and final measurement.

Typically, these models specify pupil test-scores in the current year as a function of pupil, family, school, and teacher characteristics in the current year and pupils' test-score from a previous academic year. In the literature, the model has two specifications; (1) the "gain specification" in which the lagged test-score is used to compute the dependent variable. It is used to generate the difference between current and past test-scores and; (2) the "lagged specification" in which the lagged test-score is used as an explanatory variable.

The position of the lagged score in the model has several implications. The gain-specification does not include lagged inputs as explanatory variables. Only contemporaneous inputs are used to explain the difference in test-scores from the previous year. As made clear in Todd and Wolpin (2003), for this specification to be valid, it must be assumed that; the effect of an education input is independent of the age at which it is received; and there is no "loss" or "decay" over time in the impact of inputs received in earlier periods. For example, if a pupil is exposed to a "child friendly" practice P, and this has an impact on his test-score T of an amount $\Delta T = \beta P$; β does not depend on the point in time at which the pupil is exposed to the practice; and the impact of the practice in all future periods will remain βP. If these conditions are assumed true, then gains in current achievement need only be explained by inputs in the current academic year.

A more plausible scenario is offered by the lagged specification of the value-added model. In this model the lagged test-score is used as an explanatory variable. As Todd and Wolpin (2003) demonstrate, for the model to be valid an assumption must be made that the greater the temporal distance between inputs and the current test-score, the less explanatory power they will have for the test-score. Moreover, the rate at which inputs lose explanatory power, or "decay", must be the same for all inputs. For example, the impact of having "child friendly" practice P at age A will be $\Delta T = \beta_A P$. The effect on test-scores measured at age $A+1>A$ will be $\Delta T = \beta_{A+1} P$, where the effect of the input decays over time at a rate $0 < \gamma < 1$, such that $\alpha_{A+1} = \gamma \alpha_A$. Moreover, γ must be assumed to be the same for all inputs.

An extensive body of literature has used both the gain and lagged-specifications of the value-added model to explain pupil achievements. In the USA, Clotfelter et al (2007), Aaronson et al (2007) and Kane et al (2006) have used lagged test-scores in their achievement models to examine the effects of observed characteristics of teachers on pupil test-scores. In the UK, Slater et al (2012) use lagged test-scores from primary school exams to robustly explain pupil performance in secondary school exams due to variability in teacher quality. Dearden et al (2011) use lagged test-scores to examine the variation in impact that schools in England have on pupils with different levels of prior-attainment. In India, Kingdon and Atherton (2010) make use of the gain-specification and examine the role that contract-status of teachers take in explaining test-score differences of primary school pupils. Neither specification has previously been used to examine the effects of child-friendly practices.

3.2 Estimation Strategy

The lagged specification of the value-added model is used in the analysis, as it makes less restrictive assumptions about parameter interpretations than the gain specification. Test-scores at a particular point in time are the output of a production process that cumulates the effects of all education inputs over the pupil's life-time; a vector of the 6 child-friendly practices identified in section 2.2 *(CFP)*, other routine pedagogical practices such as copy writing, recitation and dictation *(RP)*, pupil's personal characteristics and family background (PC), teacher Characteristics *(TC),* and school characteristics *(SC).* Let *(A)* denote the vector of the history of inputs / characteristics over time. Let *i* index pupils, *j* index schools, and *t* index survey rounds. A pupil's test-scores at *t* (i.e. end of round 3 data collection), can be written as a function of the entire history of inputs that he/she has been exposed to:

$$Finalscore_{ijt} = f(CFP(A)_{ij,} RP(A)_{ij,} \mathrm{P}C(A)_{ij,} TC(A)_{ij,} SC(A)_{ij}) \qquad (1)$$

To transform this relationship into a functional form, an assumption is made that the lagged test-score (i.e. baseline score in survey round 1, indexed by *t-1*) will proxy the history of effects of all

inputs, including teaching practices. Assuming that inputs are additively separable, equation (1) can be written as:

$$Finalscore_{ijt} = const + \beta_1 CFP_{ij} + \gamma Baselinescore_{ijt-1} + \beta_2 RP_{ij} + \beta_3 TC_{ij} + \beta_4 PC_{ij} + SC_j + e_{ijt} \quad (2)$$

$Finalscore$ is the z-score transformation of test-scores obtained in survey round 3, which renders pupil achievement relative to the subject average in a grade, across all schools. β_1 to β_4 are vectors of coefficients, with β_1 being the effect of child-friendly practices on round 3 final test-scores. *Baseline score* contains z-scores of the test-scores obtained during round 1 testing, transformed in the same way as final scores. The coefficient on the baseline score, *γ*, implicitly captures the rate of decay of impact of all inputs received up to the time of the baseline score. The assumption is made that the effects of all inputs decay over time at a constant rate, which is equal for all inputs. SC_j is a school fixed-effect, capturing all school characteristics (observed and unobserved) that do not vary across grades and pupils within the school. Test-scores and practices are pooled across grade 2 and grade 4 to allow for this.

To obtain bench-mark estimates of the effects of child-friendly practices, equation 2 is estimated by regressing on the *CFP* vector only. The baseline test-score is then introduced followed by routine practices and so forth. Equation 2 is estimated for reading and maths final scores separately. Within the OLS estimations standard errors are clustered at the school-level, since school effects and the error term are likely to be correlated across pupil within schools.

Since a variation of teaching practices within schools and across grades is present in the data, it is possible to estimate equation 2 using school fixed-effects. This estimation succeeds in removing all time constant unobserved school characteristics that are likely to be correlated with the application of child-friendly practices, thereby rendering a more causal estimate of their effects on test-scores.

3.3 Potential Estimation Challenges

A number of potential challenges are present in identifying the effects of child-friendly practices. These include measurement errors, multi-collinearity, and endogeneity. The implications of and (where possible) strategies to resolve each of these are discussed in turn below.

3.3.1 Measurement Error

It is quite likely that the test-scores in the data are measured with error, as tests are at best imperfect measures of a pupil's true knowledge of the subject, and scores are affected by random factors at the time of testing[12]. If we assume that errors in measurement are not correlated with the true score of the pupil then the estimate of the coefficient on test-scores will be biased downwards due to attenuation bias. Not only this, the coefficients on the teaching practices and other variables are also likely to be biased due to this measurement error and it is generally not possible to confirm the direction or size of the bias. Strategies to deal with measurement error and limiting associated biases in coefficient estimates in the literature have been to use dynamic panel data methods (Arellano & Bond (1991) and Andrabi et al, (2011)). A robustness check using these methods is not possible with the available data, due to the insufficient measures of test-scores over time.

It is worth noting that the attenuation bias introduced by measurement errors in explanatory variables can be shown to be smaller in models utilising fixed effects than in those utilising first differences (Grilliches & Hausmann (1986)). In the education context, this implies that the effects of attenuation bias are less serious in the lagged specification used in this paper than in specifications which employ first differencing, as seen in some other papers (e.g. Metzler & Woessman (2012)).

[12] It is reasonable to expect that errors in measurements of true subject knowledge are not serially-correlated over time but are more the consequence of random occurrences (Todd & Wolpin, 2003), especially since discussions in the literature generally attribute measurement errors in test-scores to "vagaries" in testing environments (eg. Jerrim & Vignoles, 2011, Andrabi et al (2011) and Clotfelter et al, 2007,). In such a situation the correlation between the lagged test-score measurement and its error results in an attenuated coefficient estimate (on the lagged test-score).

Furthermore, the test-scores are drawn from a very "low-stakes" test setting wherein examination stresses and/or performance anxiety are less likely to induce deviation of pupils' responses from their "true" knowledge[13].

Measurement error issues could also apply to the teaching practice variables themselves. In the literature, these errors have typically been dealt with by collecting longitudinal observations of teaching practices, scoring them and taking the average of scores over time (Kane et al, (2011)). Once again, the data for this paper is not suitable for the application of this method.

It is plausible that teaching practices have been measured with minimal deviation from true behaviours, owing to the non-experimental, "no-stakes" environment data was collected in. Muralidharan and Sundaraman (2010a) examined the effect of increased teaching activity on pupil test-scores within an experimental setting. Their measures of teaching activity are what this paper considers 'routine practices'. They find increased teaching activity has insignificant effects on test-scores. They suggest that this finding may be influenced by 'Hawthorne' effects: being subject to observation may cause altered activity at the time of observation, though not necessarily otherwise[14]. In contrast, Kane et al (2011) use data on teaching practices drawn from classroom observations within a non-experimental setting. They find significant effects of these practices. It is plausible that the non-experimental settings are less affected by Hawthorne-effects and measure behaviours that are more representative of true teaching practices. There is no treatment, intervention or advertised stake for participants, reducing the scope for behaviours to be primed.

[13] Since pupils' were tested in two subjects, test-score correlations within-survey-rounds-across-subjects and across-survey-rounds-within-subjects can been calculated. Each of these correlations is at least 0.7. Cronbach's alpha has also been calculated and is found to be well above the conventional 0.7 threshold for reliability for both maths and reading tests in all survey rounds. These statistics suggest that the data underlying the lagged test-score explanatory variable has been captured with a reasonable degree of reliability and consistency both within and across survey rounds.

[14] In this case, the treatment group teachers had been given background info on teaching processes, and received multiple visits from the survey team, so were likely aware of what surveyors were interested in observing. If teachers do not habitually apply the teaching practices which were observed at time of survey, this could in turn cause the finding of insignificant effects.

This paper draws on observations made within a non-experimental setting, on practices that are plausibly harder to accelerate if not habitually practiced.

3.3.2 Multi-collinearity in Child-Friendly Practices

A concern in estimation is that the variables contained in the *CFP* vector may be highly collinear. Under collinearity, coefficient estimates are likely to change erratically. Each of the six child-friendly teaching practices is recorded as a binary variable. Standard correlation coefficients perform poorly in examining the relationships between binary variables. To help judge the extent of correlation in the *CFP* vector a tetrachoric correlation matrix has been constructed in table A.4 in the appendix. Tetrachoric correlations assume that a latent normal distribution underlies each binary variable, and estimates the correlations between variables as if they were continuous. Correlation coefficients are computed by an iterative maximum likelihood estimation process akin to a bi-variate probit model. The correlations between the variables are mostly well below 0.5. This suggests that low levels of correlations underlie the observation of practices and they are unlikely to be collinear in regression models[15].

3.3.3 Child-Friendly Practices are Endogenous

Three factors are likely to mean child-friendly teaching practices are potentially endogenous; first, the characteristics of the teachers and pupils that sort into the school, second, the propensity of teachers to use the practice and third the ability of pupils to respond to the practice. The first issue is dealt with by applying school fixed-effects. The sorting of teachers and pupils into schools is not a random process and is driven by a number of factors only some of which are observable in the data. For example, as part of policy initiatives to manage pupil-teacher ratios within schools, teachers

[15] As a further check, variance inflation factors (VIF) have been calculated for each OLS regression in Table 4. The VIF provides an index that measures how much the variance of an estimated regression coefficient is increased because of collinearity. As a rule of thumb, a variable whose VIF values are greater than 10 may merit further investigation. The VIF values for each regression are well within acceptable limits, ruling out any significant multi-collinearity.

may be hired on contracts and assigned to remote areas[16]. School fixed-effects estimation removes all time-constant unobserved factors that drive the non-random sorting of teachers and pupils into schools.

Once sorted into the schools, teachers may prefer to teach the grade with higher ability pupils[17]. Once assigned to the grade, it may be the case that teachers with certain characteristics, are motivated to ensure that their pupils learn well and apply child-friendly practices. Motivation is intrinsic to a teacher's psyche/environment and is not measured in the data. It may also be the case that motivated teachers use a practice because they are privy to pupils' unobserved characteristics (such as higher innate abilities) that make them likely to respond positively to this method[18]. To reduce these remaining issues a pupil fixed-effect specification could be applied. This estimator removes all time-invariant observed and unobserved factors that do not vary within pupils ((Slater et al (2012), Clotfelter et al (2007), Aslam and Kingdon (2011)), including any factors that drive the non-random matching of teachers to schools, then matching of teachers to grades/pupils within schools and finally child-friendly practices to teacher/pupils within grades However, the data-structure does not allow such an estimation strategy as the child-friendly practices under examination are not subject-specific and hence do not vary within the pupil. Therefore the analysis though successful in removing some sources of endogeneity is unable to account for all of them.

4. Results

The first subsection centres on the lagged specification of the value-added model in explaining the impact of the 6 child-friendly practices on final test-scores. This is followed by a discussion of results from stable school environments, which is the strongest pupil-teacher match in the dataset, and

[16] Muralidharan and Sundaraman (2010b) & Kingdon and Atherton, (2010)

[17] This is only possible across grades and not within grades, as in the data there is only one classroom per grade.

[18] Table A.5 in the appendix estimates probit models for each child-friendly practice in order to examine the extent to which practices might be endogenous in the data. Multiple teacher and school characteristics are found to be significantly associated with the propensity to use the practice.

serves as a robustness check. Finally, differential impacts of child-friendly practices on high and low-scoring groups of pupils are examined.

4.1 Effects of Child-friendly Practices in the Full Sample

Table 4 shows the results for OLS and school fixed-effects specifications of the lagged value added model for all 645 schools. Coefficients and significance levels for the six practices, pupil's grade and baseline test-scores are presented[19]. Panel A presents results for reading scores, panel B presents results for maths scores.

Column 1 is the base-specification for reading scores. It presents the association between the 6 child-friendly practices and final test-scores for reading in the absence of all other inputs. The magnitude of the coefficients on the practices are quite large, ranging from -0.11 to 0.27 S.D; with the practice "contextualizing lessons" reducing final test-scores by 0.11 S.D and the practice "displaying pupils work in classrooms" increasing final test-scores by 0.27 S.D. Five of the six practices are significantly associated with final reading test-scores. The practices "displaying pupils work" and "smiling/laughing/joking with pupils" are the most highly significant and the practice "contextualizing lessons" is the most weakly significant.

Columns 2 to 4 present OLS estimates as a series of control variables are added into the specifications. Baseline test-scores are introduced in Column 2. Their inclusion substantially reduces the size of the coefficients on child-friendly practices. For example, the coefficient on "displaying pupils' work in the classroom" falls by 20 percent to 0.20 S.D. The introduction of routine teaching practices (column 3), and pupil, teacher characteristics and school resources (column 4), further

[19]As a robustness check, specifications with baseline scores represented by z-score quintiles were estimated and the same linear relationship wherein, pupils in the higher quintiles scored higher in the final tests was found. Further, no substantial differences in child-friendly practice effects between the quintile specifications and the linear z-score specifications were found. Specifications specifying baseline scores as a quadratic and cubic function were also estimated, this did not change the main results. Finally, standardizing z-scores such that performance is relative to the average in each subject within a grade across schools *within a particular state* did not change the main-results.

reduce the impact of child-friendly practices. The coefficient on "displaying children's work in the classroom" falls a further 60 percent to 0.079 S.D. when all controls are added into the specification in column 4. The significance levels of the practices are also affected as control variables are added. In column 4, only two out of the six child-friendly practices are significant, as opposed to five in the base specification in column 1. These are the practices "contextualizing lessons" and "encouraging pupils to ask questions". The latter is only weakly significant at the 10 percent level. It is important to note that even though the magnitudes of the coefficients on these practices do not fall as prominently as the practice of displaying children's work in classrooms across the specifications, the influence of these practices on final test-scores is negative and opposite to what is expected in the NCF.

The fully-specified school fixed-effects model is presented in columns 5. This is the preferred specification, in which all time-constant observed and unobserved school level characteristics that drive the sorting of pupils and teachers into schools are removed. The coefficient sizes on the practice variables are much smaller than in the base specification. For example, "displaying pupils work in the classroom" was initially associated with a 0.27 S.D increase in final reading scores, but now falls to 0.08 S.D. Each coefficient is also now statistically insignificant at conventional thresholds, indicating that there is no evidence that child-friendly practices as promoted in the NCF actually have a positive impact on pupil learning. Instead, we see from the tests of joint significance of the control variables at the bottom of Table 4 that it is the standard pupil (including family) and teacher characteristics that are significant drivers of learning[20].

Results for maths are presented in columns 6 to 10, in the same order as is used for reading. Substantive conclusions on the effects of child-friendly practices remain intact if maths test-scores

[20] Table A.3 in the appendix presents results for the full models. Routine teaching practices that significantly and positively affect reading test-scores are "pupils reading aloud from text-book". For maths "pupils writing on blackboards" and "reciting together" exert positive influences. Pupil characteristics that significantly and positively affect test-scores in both subjects include mothers' education, private tuitions and regular attendance to school. Teacher characteristics that are significant for both subjects are higher education degrees and receipt of professional in-service/pre-service teaching training. The latter is negatively associated with test-scores.

are used instead of reading. In column 6, the base specification, three out of six child-friendly practices are significantly associated with final maths scores in. The introduction of baseline test-scores (column 7), and other contemporaneous inputs (columns 8 and 9) again reduce the size and significance of the coefficients on the child-friendly practices. Within the school fixed-effects specification in Column 10, as with reading, most child-friendly practices are insignificant. The only notable difference between the school fixed-effects specifications for the two subjects is that the effect of having work displayed in the classroom is weakly significant for maths (at the 10 percent level). However, the effect of this practice on maths scores is the opposite of what is expected by policy: its presence reduces final maths test scores by 0.14 S.D. Once again it is the standard pupil and teacher characteristics that are the most significant drivers of learning in mathematics[2122].

[21] Several checks have been carried out to confirm the robustness of the results presented. Gain-specifications, two-level MLM specifications, Random Effects specifications and separate state-wise regressions with school fixed-effects of the lagged-value added model were estimated. None of these estimations yielded evidence that the six practices had consistently positive and significant effects on test-scores.

[22] Table 4 identifies the *direct* effects of each of the six practices on test-scores. As a robustness check all six practices were summed into an index "CHILDFRIENDLY". The school fixed effects specifications in Table 4 were then re-estimated with this index. No positive and significant effects of this index on test-scores were found. As a further check, (given the strong assumption behind an equal-weight summation), a second index was calculated, "CHILDFRIENDLYSCR", as the output of factor analysis of the polychoric correlation matrix of all six practices and the school fixed effects specifications in Table 4 were re-estimated. Once again, no positive and significant effects of this second index on test-scores were found

TABLE 4: OLS & School Fixed-Effects Results for the Full Sample

	PANEL A: Reading Scores					PANEL B: Maths Scores				
Dependent Variable: Final Score (standardized)	1	2	3	4	5	6	7	8	9	10
	OLS^ ~				School Fixed Effects~	OLS^ ~				School Fixed Effects~
Work Displayed in classroom = 1	0.274 (0.070)***	0.204 (0.058)***	0.181 (0.059)***	0.079 (0.049)	0.082 (0.079)	0.243 (0.073)***	0.171 (0.065)***	0.147 (0.065)**	0.027 (0.055)	-0.145 (0.086)*
Teachers Smiled/laughed/joked = 1	0.243 (0.069)***	0.19 (0.059)***	0.16 (0.060)***	-0.020 (0.050)	0.023 (0.095)	0.315 (0.073)***	0.256 (0.065)***	0.217 (0.066)***	0.024 (0.056)	-0.002 (0.009)
Teachers contextualized lessons = 1	-0.116 (0.068)*	-0.111 (0.058)*	-0.143 (0.059)**	-0.116 (0.049)**	-0.065 (0.079)	-0.11 (0.066)*	-0.141 (0.058)**	-0.175 (0.062)***	-0.11 (0.052)**	-0.156 (0.107)
Pupils encouraged to ask questions = 1	0.005 (0.069)	-0.064 (0.056)	-0.107 (0.058)*	-0.085 (0.043)*	-0.009 (0.066)	0.013 (0.07)	-0.034 (0.059)	-0.083 (0.064)	-0.066 (0.054)	0.005 (0.085)
Teachers used materials to teach other than text books = 1	0.175 (0.082)**	0.055 (0.069)	0.051 (0.071)	0.021 (0.058)	0.058 (0.108)	0.1 (0.093)	-0.001 (0.084)	-0.038 (0.082)	-0.079 (0.058)	0.103 (0.11)
Teachers organized group activities = 1	0.158 (0.068)**	0.134 (0.060)**	0.124 (0.061)**	0.048 (0.053)	0.033 (0.087)	-0.06 (0.082)	-0.031 (0.072)	-0.047 (0.073)	-0.079 (0.058)	-0.113 (0.079)
Grade 2 = 1	0.032 (0.042)	0.026 (0.037)	0.026 (0.037)	0.041 (0.043)	0.009 (0.046)	0.036 (0.045)	0.028 (0.041)	0.014 (0.042)	0.035 (0.047)	0.038 (0.047)
Baseline Test Score (Standardized)		0.398 (0.018)***	0.394 (0.018)***	0.330 (0.157)***	0.27 (0.016)***		0.349 (0.020)***	0.34 (0.020)***	0.275 (0.165)***	0.223 (0.016)***
Stable School = 1				-0.014 (0.047)					-0.105 (0.049)**	
_*cons*	-0.165 (0.044)***	-0.107 (0.038)***	-0.228 (0.055)***	-0.182 (0.362)	-1.071 (0.754)*	-0.124 (0.049)**	-0.071 (0.049)***	-0.218 (0.072)***	-0.286 (0.437)	1.532 (0.848)*
R^2	0.04	0.19	0.2	0.3	0.12	0.03	0.15	0.16	0.27	0.09
Number of Pupils	9474	9474	9474	9474	9474	9536	9536	9536	9536	9536
Routine Teaching Practices	No	No	Yes***	Yes*	Yes	No	No	Yes***	Yes	Yes**
Pupil Characteristics	No	No	No	Yes***	Yes***	No	No	No	Yes***	Yes***
Teacher Characteristics	No	No	No	Yes	Yes***	No	No	No	Yes	Yes***
School Characteristics	No	No	No	Yes***	No	No	No	No	Yes*	No
Regional Dummies	No	No	No	Yes***	No	No	No	No	Yes***	No

* p<0.1; ** p<0.05; *** p<0.01 All standard errors are in brackets, clustered at the school level and corrected for heteroscedasticity. Routine teaching practices included whether teachers wrote on blackboards, read aloud from text-books, made pupils do any kind of written work (usually 'copy-writing' from black-board), asked pupils easy oral questions, checked pupils notebooks/slates, gave dictation, made pupils recite lessons together, asked pupils to come up to the front and write on the black-board. Tetrachoric matrices showed low levels of correlation between all routine pedagogical practices (including with child-friendly practices). School Fixed-Effects specifications estimated with and without routine practices showed similar results for child-friendly practices. Pupil characteristics control for gender, age, religion, health-history, mother tongue, family SES, mothers education, whether they took paid tuitions outside of school, whether their homes had printed materials other than text-books. Teacher characteristics include teachers' age, gender, education qualifications, professional training, years of experiences, contract status. School characteristics include baseline school resources (eg. working library, toilets, TLM availability) and mid-day meal. Regional dummies include the Indian state that observations have been drawn from. Missing dummies have been included for routine teaching practices, pupil/teacher/school characteristics in the regressions.

4.2 Robustness Check: Effects of Child-friendly Practices in the "Stable" School Sample

As noted in section 2, two types of schools are identified in the data: "stable" schools, in which a single teacher was identified as holding sole responsibility for teaching reading or maths to either grade 2 or grade 4; and "flexible" schools, in which multiple teachers teach pupils a subject, but one teacher was identified as teaching pupils most frequently. The results presented in Table 4 are for the full sample, including both "stable" and "flexible" schools. For the flexible schools, the most frequent teacher's characteristics have been retained. The full OLS specifications in Table 4 (columns 4 and 9) include a dummy indicator for whether a school was stable or not. While there are no significant differences in final test-scores attained across stable and flexible schools for reading, there seem to be significant differences for maths. This dummy is assigned at a school level, so cannot be included in the school-fixed effects regressions (columns 5 and 10). However, it might be postulated that within the "flexible" schools the practices of other teachers have influenced test-scores, diluting the estimated effects of the child-friendly practices of the observed teacher. Therefore, as a robustness check, the school fixed effects specifications are re-estimated for the sub-sample of "stable" schools only.

Table 5 presents the "stable" schools only re-estimations for the school fixed-effects specifications for reading and maths. Somewhat stronger effects of child-friendly practices are found, as is evidenced by the larger magnitudes of the coefficients on the practice variables in Table 5 as compared to Table 4. For example, the coefficient on "teachers' smiling/laughing/joking" is 0.19 S.D. as compared to the 0.02 S.D school fixed effect estimate for reading in Table 4. There are also more child-friendly practices that are significant in Table 5 than Table 4 for both subjects. For example, in Table 5, two out of six practices significantly affect final reading test-scores as compared to none in Table 4. Three out of six practices significantly affect final maths scores in Table 5 as compared to one in Table 4. However, despite the increase in magnitude of the coefficients, across the subjects, four out of the five significant coefficients are very weakly significant (at 10 percent levels).

Furthermore, as a seen before in Table 4, not all of the significant practices affect learning positively within the stable school sample. This is evidenced by the negative coefficients on the practice of "organizing small group activities". This practice reduces reading and maths test-scores within the stable schools by 0.35 S.D. and 0.23 S.D. respectively.

Combining evidence from both Tables 4 and 5 it is reasonable to conclude that even though the effects of child-friendly practices show some responsiveness to a more stable teaching responsibility environment, in general, (and in contrast to what is clearly expected in policy), their effects on pupil learning are largely insignificant (or at best weakly significant at conventional thresholds), and in some cases even appear to be severely negative.

Table 5: School FE Results for "Stable" Schools

Dependent Variable: Final Score (standardized)	**1 READING**	**2 MATHS**
Work Displayed in classroom = 1	0.034 (0.120)	0.267 (0.149)*
Teachers Smiled/laughed/joked = 1	0.195 (0.114)*	0.200 (0.114)*
Teachers contextualized lessons = 1	-0.121 (0.094)	-0.107 (0.122)
Pupils encouraged to ask questions = 1	0.037 (0.073)	-0.046 (0.088)
Teachers used materials to teach other than text books = 1	0.092 (0.095)	0.163 (0.111)
Teachers organized small group activities = 1	-0.351 (0.103)***	-0.234 (0.128)*
Grade 2 = 1	-0.013 (0.065)	-0.038 (0.078)
Baseline Test Score (Standardized)	0.312 (0.020)***	0.240 (0.207)***
Const	-1.67 (0.797)**	-0.232 (1.197)
Within R^2	0.15	0.11
Number of Pupils	5,036	4,867
Routine Teaching Practice Included	Yes	Yes*
Pupil Characteristics Included	Yes***	Yes***
Teacher Characteristics Included	Yes***	Yes**

* p<0.1; ** p<0.05; *** p<0.01; Standard Errors are in brackets, clustered at school level and corrected for heteroscedasticity.
Missing dummies have been included for routine teaching practices, pupil/teacher/school characteristics.

4.3 Effects of Child-Friendly Practices on Different Ability Groups

The NCF states that "schools need to be centres that prepare children for life and ensure that all children [regardless of physical and cognitive abilities] get the maximum benefit… from an education" (*Section 4.3.2, Paragraph 2, Pg. 85*). Within the same paragraph, the document expresses criticism of "tendencies" in which small groups of pupils are routinely selected and focused on by teachers within the classroom. It is suggested that along with other factors (such as socio-economic status, caste-groups, gender etc.), these pupils are perceived as being of higher academic abilities by teachers. In view of this, this section examines whether child-friendly practices impacts groups of pupils with varying abilities differentially. Following the discussion in the NCF, it could be the case that pupils of high abilities receive more focused attention over the course of the academic year and therefore child-friendly practices interact differently with these pupils as compared to pupils of lower ability.

Ability is an innate quality of a pupil, not determined by the classroom environment. Since this quality is difficult to observe, the most appropriate marker to measure it in the dataset is pupils' baseline test scores. For each individual pupil in the dataset, an identifier has been generated to indicate whether they score in the top 20%, bottom 20% and in the middle of their grade at the time of baseline testing. This has been done to create "groups" of pupils that teachers may differentially focus child-friendly practices on. The top and bottom identifier is then interacted with each of the 6 child-friendly practices, and entered into regressions. Equation 3 below is estimated with school fixed-effects for reading and maths for the full and stable school samples and presented in Table 6.

$$Finalscore_{ijt} = const + \beta_1 CFP_{ij} + \beta_2 RP_{ij} + \beta_3 TC_{ij} + \beta_4 PC_{ij}$$

$$+\beta_5 top_{ijt-1} + \beta_6 bottom_{ijt-1} + \beta_7 (top * CFP)_{ij} + \beta_8 (bottom * CFP)_{ij}$$

$$+SC_j + e_{ijt} \qquad (3)$$

Significant differences emerge in both final reading and maths scores obtained between the group of pupils ranked as scoring at the top and bottom of the grades at baseline and pupil groups scoring in the middle within the full and stable school samples. Top-ranked pupils obtain higher final test-scores and low ranked pupils score significantly lower on the final tests than pupils ranked in the middle of the class. As for the vector of interaction effects contained in β_7 and β_8; no strong evidence emerges that the six practices interact in consistently positive or negative ways with rankings of pupil performance on the baseline tests. In the few instances that there is a significant interaction, the significance levels are generally weak. Based on the evidence, it is not possible to judge conclusively which practices benefit top/bottom ranked groups pupils more or less than others[23].

In summary, the specifications examined throughout Section 4 show that the child-friendly practices largely have insignificant effects on pupils' test-scores, and in some cases may even have a negative impact.

[23] As a robustness check, a specification with the continuous standardized baseline scores interacted with each child-friendly practice variable was estimated. The interaction effects were found to be once again, not consistently positively significant in this specification. Also, quantile regressions confirms that the six practices do not have consistent positive effects on performance of pupils in different percentiles of the test-score distribution for both subjects.

Table 6: School FE Results for "Top" & "Bottom" Scoring Pupils

	PANEL A: READING		PANEL B: MATHS	
	Full Sample	Stable Schools	Full Sample	Stable Schools
Work Displayed in classroom = 1	0.150	0.106	-0.137	0.343
	(0.095)	(0.155)	(0.106)	(0.189)*
Teachers Smiled/laughed/joked = 1	0.081	0.137	0.045	0.217
	(0.111)	(0.144)	(0.108)	(0.130)*
Teachers contextualized lessons = 1	-0.125	-0.152	-0.126	-0.109
	(0.092)	(0.132)	(0.122)	(0.160)
Pupils encouraged to ask questions = 1	0.029	0.063	0.017	-0.034
	(0.075)	(0.094)	(0.098)	(0.106)
Teachers used materials to teach other than text books = 1	0.074	0.065	0.152	0.242
	(0.115)	(0.127)	(0.127)	(0.156)
Teachers organized group activities = 1	-0.043	-0.484	-0.179	-0.312
	(0.098)	(0.126)***	(0.095)*	(0.149)*
Grade 2 = 1	0.036	-0.010	0.061	-0.024
	(0.051)	(0.075)	(0.053)	(0.092)
top' performing pupil = 1	0.287	0.339	0.279	0.290
	(0.032)***	(0.043)***	(0.036)***	(0.044)***
bottom' performing pupil = 1	-0.166	-0.269	-0.161	-0.181
	(0.032)***	(0.048)***	(0.029)***	(0.043)***
top pupil*displaying work	-0.043	-0.053	0.034	0.024
	(0.056)	(0.072)	(0.055)	(0.080)
bottom pupil*displaying work	-0.103	-0.061	-0.086	-0.073
	(0.065)	(0.074)	(0.058)	(0.080)
top pupil*friendly teacher	0.045	0.042	0.001	0.020
	(0.058)	(0.077)	(0.057)	(0.072)
bottom pupil*friendly teacher	-0.115	-0.171	-0.075	-0.118
	(0.067)	(0.073)**	(0.062)	(0.079)
top pupil*teacher contextualizing	-0.023	-0.026	-0.086	-0.064
	(0.056)	(0.088)	(0.054)*	(0.080)
bottom pupil*teacher contextualizing	0.055	0.017	-0.016	0.005
	(0.058)	(0.078)	(0.059)	(0.079)
top pupil*pupils asking questions	-0.048	0.011	-0.099	-0.065
	(0.058)	(0.075)	(0.055)*	(0.077)
bottom pupil*pupils asking questions	0.015	0.056	-0.017	0.019
	(0.059)	(0.079)	(0.055)	(0.080)

	Table 6	***(continued)***		
top pupil*teachers using TLMs	0.012 (0.063)	-0.102 (0.085)	-0.017 (0.065)	-0.37 (0.095)
bottom pupil*teachers using TLMs	0.013 (0.078)	0.102 (0.101)	0.024 (0.075)	-0.070 (0.093)
top pupil*group activity	0.022 (0.053)	-0.022 (0.080)	-0.046 (0.059)	0.009 (0.095)
bottom pupil*group activity	0.076 (0.072)	-0.020 (0.091)	0.117 (0.067)*	0.078 (0.099)
_cons	-1.87 (0.963)*	-3.48 (0.974)***	1.15 (0.904)	-0.799 (1.31)
Within R^2	0.08	0.12	0.07	0.09
Number of Pupils	9226	4,840	9291	4,688
Routine Teaching Practice Included	Yes*	Yes*	Yes**	Yes
Pupil Characteristics Included	Yes***	Yes***	Yes***	Yes***
Teacher Characteristics Included	Yes**	Yes***	Yes**	Yes*

*p<0.1; **p<0.05; ***p<0.01

Standard errors are in brackets, clustered at school level and corrected for heteroscedasticity.

Missing dummies included for routine practices, pupil and teaching characteristics.

Reference category in these regressions "pupils scoring in the middle of the class"

5. Conclusion

The idea of child-friendly practices has gained increasing currency in Indian education policy in recent years. The National Curriculum Framework (2005) strongly recommends the use of child-friendly pedagogical practices such as increasing pupil participation during classroom sessions, shifting away from rigid classroom teaching structures, creating more lively classrooms and drawing on pupils' experiences to enrich lessons and provide examples. The Right to Education Act (2009) re-emphasizes the use of these practices. Teacher training courses within the country continue to incorporate aspects of child-friendliness within their curriculums despite an apparent absence of a quantitative evidence base that the practices have consistent and positive effects on learning. The

explicit expectation of policymakers is that these practices will positively impact on learning, and indeed, these practices likely appear attractive to most readers of this paper. However, despite the increasing faith they place in such practices, nowhere in official documents is evidence given of their effectiveness.

This paper provides the first quantitative examination on the effectiveness of child-friendly practices in improving pupil learning. Six practices are analysed, each representing a different aspect of child-friendliness within the NCF. Data on the incidence of these practices is drawn from a survey of grade 2 and grade 4 classrooms within 645 government-run primary schools in rural India. Estimates of their impact on "low-stakes" reading and maths tests are generated using a school-fixed effects value-added model. The research methodology succeeds in removing any biases at the school level relating to the manner in which teachers (and their propensity to use these practices) match to pupils. Teaching practice data is drawn from a "non-experimental" setting, so the results can be expected to be largely free of any biases caused by Hawthorne effects.

The main findings that emerge are; (1) Child-friendly practices largely have insignificant effects. Pupil, family and teacher characteristics are found to be much more significant drivers of learning. (2) Even in situations where Child-Friendly practices show some effects ("stable" teaching environments where pupils are taught by the same subject teacher for the entire academic year), these are not always positive and (3) No strong evidence emerges that Child-Friendly practices differentially affect high or low ability pupils.

The implication of these findings is not that child-friendly practices should be abandoned. Though, as it currently stands, there is no strong evidence to support that child-friendly practices work. It may well be a case of "getting the basics right" before focusing on teachers behavioural traits. Inputs such as lower rates of teacher absenteeism or more stable teaching responsibility setup within schools may well be required before child-friendly practices become significant enablers of learning. At base, it would be highly beneficial if a clearer articulation was made at the point of

policy formulation of the intended causal linkage between policy input(s) and learning outcomes. Further, more generally speaking, policy should be formed on the basis of a strong and systematic evidence base as to what works. So, as the focus on introducing these practices within schools increases, it is equally important to design mechanisms that routinely evaluate and quantitatively track their effects on learning outcomes.

Further research in this area could profitably be focused on examining the impact of child-friendly practices in a wider range of contexts and outcomes than is possible here. In particular, evidence of their effectiveness in less deprived schools (eg. urban schools in India or even within more developed country's schools such as the UK) and non-cognitive outcomes would be of interest. It would also be highly valuable to expand into psychological evaluation of the effects and effectiveness of these practices, before they are further emphasised in policy.

References

ASER, 2005-2010, 'Annual Status of Education Report (Rural)' *Pratham Resource Centre, available at*, http://www.asercentre.org/ngo-education-india.php?p=Download+ASER+reports.

Aaronson, D., Barrow, L., & Sander, W. (2007). 'Teachers and student achievement in the Chicago public high schools'. *Journal of Labor Economics*,*25*(1), 95-135.

Andrabi, T., Das, J., Khwaja, A. I., & Zajonc, T. (2011). 'Do value-Added estimates add value? Accounting for learning dynamics'. *American Economic Journal: Applied Economics*, *3*(3), 29-54.

Arellano, M., & Bond, S. (1991). 'Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations'. The Review of Economic Studies, 58(2), 277-297.

Aslam, M., & Kingdon, G. (2011), 'What can teachers do to raise pupil achievement?'. *Economics of Education Review*, *30*(3), 559-574.

Atherton, P., & Kingdon, G. (2010). 'The relative effectiveness and costs of contract and regular teachers in India'. *Centre for the Study of African Economies (CSAE) Working Paper Series*, *15*.

Banerji, R. (2000). 'Poverty and primary schooling: Field studies from Mumbai and Delhi'. *Economic and Political Weekly*, 795-802.

Bhattacharjea, S., W. Wadhwa, R. Banerji, 2011, 'Inside Primary Schools: Teaching and Learning in Rural India', *Pratham Mumbai Education Initiative, available at* http://images2.asercentre.org/homepage/tl_study_print_ready_version_oct_7_2011.pdf.

Clotfelter, C., H. Ladd, J. Vigdor, 2006. 'Teacher-Student Matching and the Assessment of Teacher Effectiveness', *No. w11936, National Bureau of Economic Research*

Dearden, L., Micklewright, J., & Vignoles, A. (2011). 'The Effectiveness of English Secondary Schools for Pupils of Different Ability Levels*'. *Fiscal Studies,* 32(2), 225-244.

Duflo, E., P. Dupas, M. Kremer, 2009. 'Additional Resources Versus Organizational Changes in Education: Experimental Evidence from Kenya'. *Unpublished manuscript. Abdul Latif Jameel Poverty Action Lab (JPAL), Cambridge, Mass.: Massachusetts Institute of Technology*

Glewwe, P., N. Illias, M. Kremer, 2003. 'Teacher Incentives'. *No. w9671. National Bureau of Economic Research*

Griliches, Z., & Hausman, J. A. (1986). 'Errors in variables in panel data'. *Journal of econometrics*, *31*(1), 93-118.

Hanushek, E. A. (2003). 'The Failure of Input-based Schooling Policies*'. *The economic journal*, *113*(485), F64-F98.

Jerrim, J., & Vignoles, A. (2011). 'The use (and misuse) of statistics in understanding social mobility: regression to the mean and the cognitive development of high ability children from disadvantaged homes'. *(No. 11-01). Department of Quantitative Social Science-Institute of Education, University of London*.

Kane, T. J. and Staiger, D. O. (2002). 'The promise and pitfalls of using imprecise school accountability measures'. *Journal of Economic Perspectives,* 16 (4)*,* 91 - 114.

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). 'What does certification tell us about teacher effectiveness? Evidence from New York City'. *Economics of Education Review*, *27*(6), 615-631.

Kane, T. J., Taylor, E. S., Tyler, J. H., & Wooten, A. L. (2011). 'Identifying effective classroom practices using student achievement data'. *Journal of Human Resources*, 46(3), 587-613.

Kingdon, G. G. (2006). 'Teacher characteristics and student performance in India: A pupil fixed effects approach'.

Kingdon, G., R. Banerji and P. Chaudhary (2008) "SchoolTELLS Survey of Rural Primary Schools in Bihar and Uttar Pradesh, 2007-08". *Institute of Education, University of London.*

Kremer, M., Chaudhury, N., Rogers, F. H., Muralidharan, K., & Hammer, J. (2005). 'Teacher absence in India: A snapshot.' *Journal of the European Economic Association*, *3*(2-3), 658-667.

Lavy, V. (2011). 'What makes an effective teacher? Quasi-experimental evidence'. *(No. w16885). National Bureau of Economic Research.*

Metzler, J., & Woessmann, L. (2012). 'The impact of teacher subject knowledge on student achievement: Evidence from within-teacher within-student variation'. *Journal of Development Economics*, 99(2), 486-496.

Muralidharan, K., & Sundararaman, V. (2009). 'Teacher performance pay: Experimental evidence from India'. *(No. w15323). National Bureau of Economic Research*.

Muralidharan, K., & Sundararaman, V. (2010a). 'The Impact of Diagnostic Feedback to Teachers on Student Learning: Experimental Evidence from India*'.*The Economic Journal*, 120(546), F187-F203.

Muralidharan, K., & Sundararaman, V. (2010b). 'Contract Teachers: Experimental Evidence from India'. *UC San Diego*.

National Council of Education Research and Training (2005). National Curriculum Framework. *New Delhi*

Pritchett, L., & Beatty, A. (2012). 'The negative consequences of overambitious curricula in developing countries'. *Center for Global Development Working Paper*, (293).

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). 'Teachers, schools, and academic achievement'. *Econometrica*, *73*(2), 417-458.

Slater, H., Davies, N. M., & Burgess, S. (2012). 'Do Teachers Matter? Measuring the Variation in Teacher Effectiveness in England*'. *Oxford Bulletin of Economics and Statistics*, *74*(5), 629-645.

Todd, P. E., & Wolpin, K. I. (2003). 'On the specification and estimation of the production function for cognitive achievement*'. *The Economic Journal*, *113* (485), F3-F33.

UNICEF'S, C. F. S. (2010). 'Child-Friendly Schools'. *available at* http://www.unicef.org/publications/files/Child_Friendly_Schools_Manual_EN_040809.pdf

Appendix

Table A.1: Description of Surveys

Survey Unit	Round	Key Information
Pupil	Round 1 Round 3	Personal Characteristics, Reading & Maths tests at start and end of academic year
School	Round 1 Round 2 Round 3	Pupil enrolment & attendance rates, Resources & Infrastructure observations Teacher responsibilities, appointment changes & attendance rates (from head-teacher)
Teacher	Round 1	Personal characteristics, Education, training, experience, income Responsibilities at the school
Classroom Observations	Round 1	Learning resources available in the classroom Teaching methods used (based on 30 minute observations grade 2 & 4 teachers) Student activities (based on 30 minute observations of grade 2 & 4 pupils) Use of teaching & learning resources by both pupils & teachers
Household	Round 2	Family characteristics for sampled pupils (personal & education) Household economic characteristics Home literacy environment Details of academic support available to the sampled pupil Health history of the sampled pupil

Table A.2: Sample Descriptions

This table compares the average characteristics of the sub-samples used in the paper with the main IPS survey sample. It is encouraging to see that there are hardly any significant differences in teaching practices, teacher and school characteristics in the flexible + stable school sample used to generate Table 4 in the main paper. Significant differences in only 4 out of 20 pupils' characteristics are seen. The stable school samples show no differences in child-friendly practices. 3 out of 8 routine practices, 10 out of 20 pupil characteristics, 1 out of 7 teacher characteristics and 3 out of 5 school characteristics have averages that are significantly different from the full IPS sample. Since the teaching responsibility environments are different within "stable" schools It is reasonable to expect some differences in the average characteristics.

	MAIN IPS Sample	Full (Flexible + Stable School) Sample	Stable School Sample
Child Friendly Practices			
Classrooms with pupils work displayed	0.26	0.23 **	0.25
Classrooms where teachers smiled/laugh	0.23	0.23	0.26
Classrooms where teachers contextualized	0.21	0.22	0.22
Classrooms where pupils asked questions	0.28	0.27	0.30
Classrooms where teachers used TLMs	0.14	0.13	0.16
Classrooms with small group activity	0.17	0.16	0.19
Routine Practices			
Teacher wrote on blackboard	0.67	0.67	0.68
Teacher read aloud from textbook	0.64	0.65	0.68**
Pupils doing any kind of written work	0.63	0.62	0.62
Teachers asked pupils oral questions	0.57	0.58	0.63
Teachers checked written work	0.51	0.51	0.54
Teachers gave dictation	0.39	0.39	0.41**
Pupils were made to recite singly/together	0.37	0.37	0.43**
Pupils asked to write on blackboards	0.26	0.26	0.29
Pupils			
Baseline Reading (Grade 2)	13.0	12.7	13.8 **
Final Reading (Grade 2)	17.5	17.0	18.3 **
Baseline Maths (Grade 2)	6.6	6.6	6.8
Final Maths (Grade 2)	7.9	7.8	8.1**
Baseline Reading (Grade 4)	3.4	3.5 **	3.8**
Final Reading (Grade 4)	4.6	4.6	5.0 **
Baseline Maths (Grade 4)	9.0	9.3 **	10.1 **
Final Maths (Grade 4)	10.6	10.6	11.4
Female pupils	0.50	0.50	0.51
Avg Age	8.2	8.3	8.1
Avg years of mothers education	2.1	2.0 **	2.4 **
Attended a pre-school program	0.45	0.43	0.52 **
Taking private tuition	0.12	0.12	0.09 **
living in permanent houses	0.32	0.30 **	0.34
Avg no. of family members	5.8	5.9	5.6 **
Pupils home have print materials (0-6)	1.7	1.7	1.7
Pupil is Hindu	0.82	0.82	0.83

Table A.2 (*continued)*

Pupil ill in the last 6 mths	1.8	1.8	1.8
Pupil found at school in all 3 survey rounds	2.1	2.1	2.2
Durable ASSET index (0-4)	1.8	1.8	1.9 **
Teachers			
Avg Age	39.2	39.1	39.0
Male teachers	0.66	0.66	0.64
Contract teachers	0.21	0.22	0.17 **
Avg. years of experience at current school	5.0	5.1	4.0
Had a higher degree	0.56	0.56	0.57
Teacher income (log)	8.8	8.8	8.8
Had some professional training	0.88	0.88	0.88
Schools			
Useable Toilets	0.33	0.31	0.35
Boundary Walls	0.27	0.28	0.24
Working Drinking water facilities	0.38	0.36	0.39
Useable libraries	0.24	0.24	0.29 **
Free mid-day meals	0.86	0.87	0.90 **
Average number of pupils enrolled	41.6	40.3	33.3 **

** indicates significant differences from the main IPS sample below the 5 percent level

Table A.3: Variables used for analysis (T-tests of difference: by Child-Friendly practices)

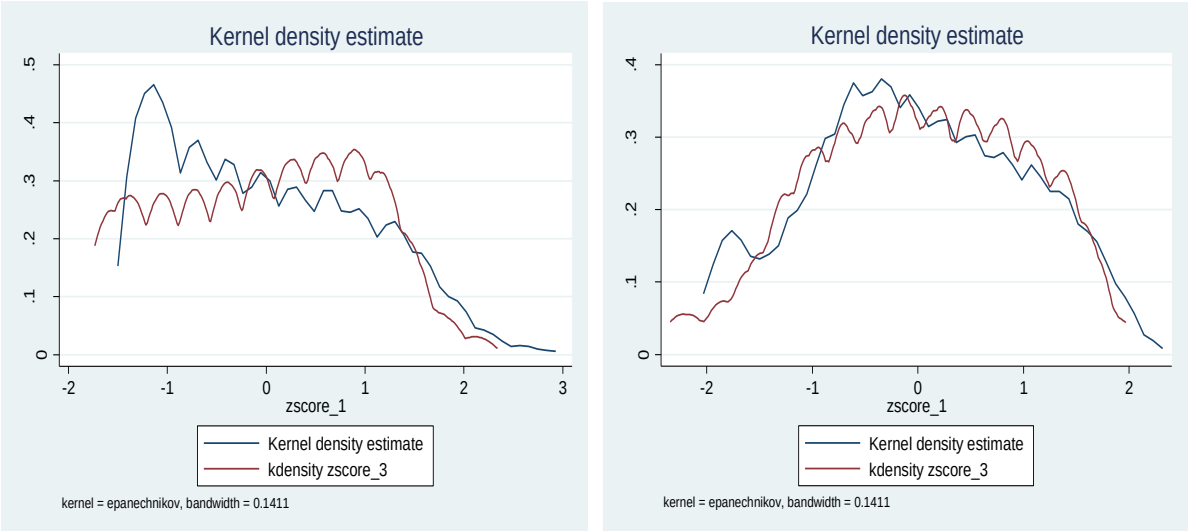
Descriptions of variables used in analysis are presented; categorized by whether or not at least one child friendly practice was observed during classroom observations. A larger proportion of pupils in classrooms with at least one child-friendly practice have mothers with more years of schooling, attended pre-schools, live in houses constructed of higher quality materials, come from smaller sized families and have slightly lower reports of falling ill in the 6mths prior to the household survey. There is no significant difference in years of experience or income levels between teachers that used a child-friendly practice and those that did not. Schools that had classrooms with child-friendly practices were significantly more likely to have a working library and a free mid-day meal (in table).

	GRADE 2			GRADE 4		
	No Practice	With at least one practice	Diff.	No Practice	With at least one practice	Diff.
Pupils						
Female pupils	0.5	0.51	-0.01	0.5	0.5	0
Avg Age	7.3	7.2	0.1	9.3	9.2	0.1*
Avg years of mothers education	1.7	2.2	-0.5***	2.0	2.2	-0.2***
Attended a pre-school program	0.37	0.51	-0.14***	0.41	0.52	-0.11***
Taking private tuition	0.13	0.08	0.05**	0.14	0.11	0.1
living in permanent houses	0.3	0.28	0.02**	0.3	0.34	-0.04**
Avg no. of family members	6.2	5.8	0.4*	6.0	5.7	0.3***
Pupils homes have print materials (0 – 6)	1.8	1.9	0.1*	2	2	0
Pupils home language matched school language	0.78	0.74	0.04**	0.74	0.73	0.1
Pupil is Hindu	0.72	0.7	0.02	0.7	0.73	-0.03**
Pupil ill in the last 6 mths	0.1	0.2	0.1**	0.1	0.2	0.1**
Pupil found at school during all 3 rounds of data-collection	2.1	2.1	0	2.1	2.2	-0.1**
Durable Asset Index (0 – 4)	1.7	1.8	-0.1**	1.7	2	-0.03**
Teachers						
Avg Age	37.5	38.8	-1.3	40.2	38.8	1.4
Male teachers	0.66	0.67	-0.01	0.74	0.67	0.07
Contract teachers	0.24	0.31	0.07	0.26	0.16	0.10**
Avg. years of teaching experience	5.6	5.3	0.3	6.0	5.5	0.5
Had a higher degree	0.41	0.52	0.11**	0.41	0.57	-0.16***
Teacher income (log)	8.7	8.7	0	8.6	8.7	-0.1
Had some professional training	0.87	0.84	0.03	0.79	0.89	-0.10**
Schools						
Available & Useable Toilets	0.45	0.52	-0.07	0.39	0.50	-0.10
Boundary Walls	0.32	0.28	0.04	0.23	0.34	-0.11**
Working Drinking water facilities	0.65	0.69	0.01	0.69	0.68	0.01
Available & Useable libraries	0.12	0.24	-0.13***	0.11	0.28	-0.17***
Free mid-day meals	0.80	0.89	-0.09**	0.81	0.9	-0.09**
Average number of pupils enrolled	45	46	-1.00	48	44	5

***p < 0.01 **p < 0.05 *p<0.1; these statistics are generated for the full-sample (i.e. pooling the "flexible" and "stable" schools). For the t-tests, standard errors are clustered at the school level

Figure 1: Pooled Test-score Distribution across Grades for Reading (left) and Maths (right)

Figure 1 shows the pooled kernel density distributions of standardized test-scores across Grade 2 and 4. The height of the kernel density shows the frequency of pupils who are at a particular standard deviation above or below the mean. There are fewer observations of pupils in the final survey round with scores of less than the group average than in the base-line, indicating positive learning across survey rounds. Performance in maths appear to be more normally distributed than in reading as evidenced by the clustering of pupils around the mean on the right of Figure 1.



Note: The blue line is the standardized scores from round 1 and the red line is the standardized scores from round 3.

Table A.4: Tetrachoric Correlation Matrix of Child-Friendly Practices (Pooled for Grade 2 and 4)

	Work Displayed in Classroom	Teachers Smiled/laughed/joked	Teachers contextualized lessons	Pupils encouraged to ask questions	Teachers used materials to teach other than text books	Teachers organized group activities
Work Displayed in Classroom	1.00					
Teachers Smiled/laughed/joked	0.27	1.00				
Teachers contextualized lessons	0.16	0.52	1.00			
Pupils encouraged to ask questions	0.18	0.42	0.36	1.00		
Teachers used materials to teach other than text books	0.40	0.18	0.26	0.22	1.00	
Teachers organized group activities	0.12	0.27	0.30	0.18	0.49	1.00

Note: Coefficients were not substantially different when matrix was generated for each grade separately.

Table A.5: Probit Models to Demonstrate the Propensity to use each Child-Friendly Practice

The dependent variable in each column of this table is a binary indicator which takes value 1 if a child-friendly practice was observed and 0 if not observed. Each column examines the propensity to use child-friendly practices. We see that the use of child-friendly practices depends mostly on the existence of other routine teaching practices within the classroom. In some cases eg. "small group activity" the use of child-friendly practices depends significantly on pupils' baseline test-scores. In general, this table displays evidence that the propensity to use a particular practice depends on a number of observed pupils, teacher and school characteristics.

	Displaying Work = 1	Teachers Smile/Joke = 1	Teachers Contextualize= 1	Pupils ask questions = 1	TLMs used = 1	Small Group Activity organized = 1
subject	0.029	-0.012	0.018	0.007	-0.002	-0.005
	(0.017)*	(0.007)	(0.013)	(0.017)	(0.009)	(0.051)
grade2 = 1	0.020	0.075	0.034	-0.083	0.009	-0.498
	(0.065)	(0.053)	(0.058)	(0.072)	(0.043)	(0.252)**
Basline Scr	0.025	0.008	0.012	-0.000	0.023	0.076
	(0.017)	(0.010)	(0.014)	(0.014)	(0.010)**	(0.048)
RoutinePractice1=1	0.007	0.065	0.189	0.058	-0.022	-0.427
	(0.080)	(0.044)	(0.044)***	(0.067)	(0.052)	(0.285)
RoutinePractice2=1	0.008	0.151	0.075	0.034	0.025	-0.440
	(0.075)	(0.032)***	(0.049)	(0.067)	(0.036)	(0.287)
RoutinePractice3=1	0.096	0.041	-0.083	0.112	0.043	0.153
	(0.062)	(0.042)	(0.058)	(0.058)*	(0.034)	(0.248)
RoutinePractice4=1	0.127	-0.080	0.074	0.160	0.051	0.186
	(0.060)**	(0.040)**	(0.045)	(0.056)***	(0.031)	(0.232)
RoutinePractice5=1	-0.027	0.099	0.067	0.123	0.062	0.293
	(0.069)	(0.043)**	(0.059)	(0.059)**	(0.039)	(0.263)
RoutinePractice6=1	0.066	-0.040	0.068	0.172	0.004	0.205
	(0.063)	(0.040)	(0.060)	(0.071)**	(0.036)	(0.225)
RoutinePractice7=1	0.128	0.063	0.099	0.151	0.074	-0.111
	(0.065)*	(0.043)	(0.051)*	(0.073)**	(0.039)*	(0.222)
RoutinePractice8=1	0.208	0.148	0.173	0.189	0.069	0.124
	(0.069)***	(0.053)***	(0.062)***	(0.072)***	(0.046)	(0.231)
Female pupil = 1	-0.007	0.010	-0.009	0.015	-0.001	0.016
	(0.021)	(0.012)	(0.016)	(0.023)	(0.012)	(0.079)
Age of Pupil	0.023	0.051	-0.000	0.003	0.009	0.009
	(0.020)	(0.018)***	(0.015)	(0.021)	(0.013)	(0.084)
Yrs. of Mthedu	0.003	0.001	0.001	-0.008	-0.002	-0.031
	(0.004)	(0.003)	(0.004)	(0.004)**	(0.002)	(0.015)**
Perm. House=1	0.007	-0.018	-0.006	-0.067	0.023	0.251
	(0.031)	(0.019)	(0.025)	(0.029)**	(0.016)	(0.105)**
Family Size	0.004	-0.004	-0.008	0.007	-0.004	-0.024
	(0.007)	(0.005)	(0.007)	(0.006)	(0.005)	(0.027)
Pre-school = 1	0.014	-0.031	-0.105	-0.035	0.022	0.045
	(0.044)	(0.025)	(0.037)***	(0.042)	(0.022)	(0.153)
Tuition = 1	0.058	-0.016	-0.035	0.036	0.028	0.428
	(0.068)	(0.050)	(0.054)	(0.069)	(0.037)	(0.187)**
Print Materials	-0.021	0.005	0.002	0.035	0.000	0.017
	(0.018)	(0.010)	(0.013)	(0.012)***	(0.010)	(0.060)
Language match	0.104	-0.128	-0.164	-0.141	0.041	-0.106
	(0.059)*	(0.055)**	(0.086)*	(0.095)	(0.034)	(0.244)
Hindu Pupil	-0.001	0.001	-0.000	-0.000	0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
Illness = 1	0.033	-0.033	-0.030	0.059	0.010	0.199
	(0.045)	(0.024)	(0.032)	(0.037)	(0.029)	(0.135)
Attendance	-0.016	-0.000	-0.018	0.040	0.005	0.112
	(0.024)	(0.015)	(0.020)	(0.023)*	(0.013)	(0.084)
Econ. Index	0.018	0.004	-0.023	-0.035	0.000	-0.089
	(0.018)	(0.010)	(0.013)*	(0.015)**	(0.008)	(0.053)*
Teacher age	0.000	0.001	-0.001	-0.001	0.002	0.018
	(0.005)	(0.003)	(0.004)	(0.004)	(0.002)	(0.015)

Table A.5 (*continued)*

Male Teacher=1	-0.038	0.069	0.015	0.004	-0.034	0.125
	(0.062)	(0.037)*	(0.049)	(0.058)	(0.036)	(0.215)
Higher degree=1	0.014	0.004	0.022	-0.061	0.032	-0.357
	(0.074)	(0.047)	(0.062)	(0.066)	(0.042)	(0.225)
Prof. Training=1	0.127	0.063	-0.030	-0.026	0.061	0.135
	(0.076)*	(0.039)*	(0.106)	(0.087)	(0.032)*	(0.351)
Yrs. of Exp.	-0.002	0.005	-0.012	-0.005	-0.007	-0.049
	(0.007)	(0.005)	(0.007)*	(0.007)	(0.004)	(0.027)*
Contract teach=1	0.081	0.074	0.303	-0.012	0.025	0.536
	(0.153)	(0.112)	(0.183)*	(0.117)	(0.080)	(0.454)
Log Tch. Income	0.091	0.036	0.170	-0.027	0.069	0.233
	(0.092)	(0.066)	(0.080)**	(0.084)	(0.043)	(0.342)
Working Toilet=1	-0.016	-0.024	0.020	0.084	0.028	-0.102
	(0.064)	(0.039)	(0.053)	(0.057)	(0.037)	(0.205)
Drinking water=1	0.018	0.023	0.006	0.041	0.018	0.489
	(0.070)	(0.039)	(0.050)	(0.059)	(0.035)	(0.228)**
Mid Day Meal=1	0.169	-0.062	0.171	-0.411	0.073	0.101
	(0.083)**	(0.120)	(0.037)***	(0.161)**	(0.041)*	(0.482)
Avg Enrolment	0.002	-0.001	0.003	0.003	0.001	-0.001
	(0.002)	(0.001)	(0.002)*	(0.001)**	(0.001)	(0.006)
Wrking Library=1	0.127	0.141	0.077	0.143	0.178	0.824
	(0.076)*	(0.056)**	(0.057)	(0.081)*	(0.064)***	(0.222)***
Boundary Wall=1	-0.041	0.011	0.053	0.049	0.089	0.253
	(0.073)	(0.043)	(0.056)	(0.070)	(0.043)**	(0.243)
Himachal = 1	-0.272	-0.170	-0.176	-0.103	-0.176	-0.041
	(0.067)***	(0.037)***	(0.050)***	(0.083)	(0.028)***	(0.530)
Rajasthan=1	0.266	-0.030		-0.184	0.066	0.656
	(0.229)	(0.080)		(0.056)***	(0.169)	(0.665)
Assam = 1	-0.181	-0.212	-0.188	-0.150	-0.081	-1.097
	(0.101)*	(0.044)***	(0.065)***	(0.085)*	(0.046)*	(0.440)**
Jharkhand = 1	-0.137	-0.079	-0.150	-0.042	-0.186	0.401
	(0.095)	(0.049)	(0.067)**	(0.084)	(0.047)***	(0.355)

*p<0.1; **p<0.05; ***p<0.01

Table A.6: Results for Full Models

This table presents the results for the fully specified model in Equation 2, estimated by both OLS and school FE. It specifically presents coefficients for all the covariates contained within columns 4 and 5 in PANEL A: READING and columns 9 and 10 in PANEL B: MATHS from Table 4 in the main paper.

	READING		MATHS	
	1	2	3	4
	OLS	School FE	OLS	School FE
Work Displayed in classroom = 1	0.080	0.082	0.027	-0.145
	(0.049)	(0.079)	(0.055)	(0.086)*
Teachers Smiled/laughed/joked = 1	-0.020	0.023	0.024	0.002
	(0.050)	(0.093)	(0.057)	(0.090)
Teachers contextualized lessons = 1	-0.117	-0.065	-0.113	-0.156
	(0.049)**	(0.079)	(0.052)**	(0.107)
Pupils encouraged to ask questions = 1	-0.086	-0.009	-0.066	0.005
	(0.043)**	(0.066)	(0.054)	(0.085)
Teachers used TLMS to teach other than text books = 1	0.021	0.058	-0.076	0.103
	(0.058)	(0.108)	(0.066)	(0.110)
Teachers organized group activities = 1	0.048	-0.033	-0.079	-0.113
	(0.053)	(0.087)	(0.058)	(0.079)
Grade 2 = 1	0.041	0.009	0.035	0.038
	(0.043)	(0.046)	(0.047)	(0.047)
Baseline Test-Score (Standardized)	0.330	0.270	0.275	0.223
	(0.016)***	(0.016)***	(0.017)***	(0.016)***
Teacher Wrote on Blackboard = 1	0.023	0.041	0.056	0.218
	(0.046)	(0.070)	(0.050)	(0.079)***
Teacher Read aloud from Text book =1	0.071	0.114	-0.018	0.088
	(0.043)	(0.057)**	(0.048)	(0.073)
Pupils doing any kind of written work = 1	-0.075	-0.092	-0.006	-0.059
	(0.043)*	(0.065)	(0.045)	(0.080)
Pupils were asked oral questions = 1	0.044	0.012	0.002	-0.034
	(0.043)	(0.067)	(0.046)	(0.071)
Teachers checked written work = 1	0.010	0.123	0.033	0.068
	(0.046)	(0.082)	(0.048)	(0.098)
Teachers gave dictation = 1	-0.078	-0.052	0.021	-0.015
	(0.042)*	(0.068)	(0.045)	(0.061)
Teachers made pupil recite singly/together = 1	-0.030	0.099	-0.024	0.189
	(0.046)	(0.082)	(0.048)	(0.073)**
Pupils asked to write on blackboard = 1	-0.004	0.006	-0.073	-0.033
	(0.049)	(0.069)	(0.052)	(0.077)
Stable School = 1	-0.014	NA	-0.106	NA
	(0.047)	NA	(0.050)**	NA
Female Pupil = 1	-0.020	-0.022	-0.112	-0.109
	(0.021)	(0.019)	(0.022)***	(0.019)***
Age of Pupils	0.000	0.006	0.001	0.012
	(0.014)	(0.012)	(0.014)	(0.011)
Yrs. of Mother Education	0.022	0.022	0.016	0.016
	(0.004)***	(0.004)***	(0.004)***	(0.004)***
Pupils live in permanent houses	-0.058	-0.005	-0.001	-0.014
	(0.030)*	(0.024)	(0.031)	(0.023)
Family size	0.006	0.005	-0.001	0.001
	(0.006)	(0.005)	(0.006)	(0.005)
Pupil attended pre-school = 1	0.056	0.010	0.031	0.002
	(0.033)*	(0.027)	(0.036)	(0.028)
Pupils takes paid tuitions = 1	0.161	0.192	0.091	0.199
	(0.041)***	(0.041)***	(0.053)*	(0.035)***

Table A.6 (*continued)*

No. of Print Materials in the household	0.033	0.028	0.027	0.016
	(0.013)**	(0.010)***	(0.013)**	(0.009)*
Pupils Home Lang matched School Lang = 1	0.017	0.042	0.167	0.051
	(0.061)	(0.058)	(0.060)***	(0.051)
Pupil is Hindu = 1	0.072	-0.015	0.012	-0.017
	(0.048)	(0.038)	(0.050)	(0.036)
Pupil ill in the last 6 mths = 1	-0.044	-0.004	-0.073	-0.029
	(0.037)	(0.029)	(0.033)**	(0.027)
Pupil found attending school in all 3 visits = 1	0.114	0.110	0.101	0.091
	(0.019)***	(0.015)***	(0.021)***	(0.015)***
Durable Asset Index (0 – 4)	-0.001	0.023	-0.005	0.026
	(0.016)	(0.013)*	(0.017)	(0.012)**
Teachers Age	0.000	0.003	-0.003	-0.000
	(0.003)	(0.004)	(0.003)	(0.006)
Male Teacher = 1	0.018	0.106	-0.002	0.144
	(0.041)	(0.072)	(0.042)	(0.084)*
Teacher had a higher degree = 1	0.013	0.167	-0.012	-0.042
	(0.042)	(0.070)**	(0.047)	(0.082)
Teacher attended a prof.l training program = 1	-0.056	-0.335	0.011	-0.431
	(0.058)	(0.130)**	(0.072)	(0.149)***
Yrs. of Teaching Experience	0.007	0.011	-0.002	0.009
	(0.005)	(0.007)	(0.006)	(0.008)
Contract Teacher = 1	0.017	-0.017	0.034	-0.098
	(0.072)	(0.106)	(0.074)	(0.156)
Log of Teacher Income	0.001	0.036	0.038	-0.218
	(0.037)	(0.085)	(0.039)	(0.095)**
School had working toilet = 1	0.047		0.051	
	(0.046)		(0.046)	
School had useable drinking water = 1	-0.062		-0.097	
	(0.044)		(0.048)**	
Mid-Day Meal available = 1	-0.145		-0.125	
	(0.059)**		(0.070)*	
Enrolment Size of School	-0.001		-0.001	
	(0.001)		(0.001)	
School had working library = 1	0.125		0.081	
	(0.053)**		(0.066)	
School had boundary wall = 1	0.055		0.029	
	(0.047)		(0.050)	
Rajasthan (ref. region HP)	-0.328		-0.333	
	(0.087)***		(0.086)***	
Assam (ref. region HP)	-0.126		0.138	
	(0.105)		(0.120)	
Jharkhand (ref region HP)	-0.562		-0.424	
	(0.082)***		(0.085)***	
Andhra Pradesh (ref region HP)	0.149		0.412	
	(0.076)*		(0.081)***	
_cons	-0.182	-1.071	-0.287	1.532
	(0.363)	(0.754)	(0.368)	(0.848)*
R^2	0.30	0.12	0.27	0.09
Number of Pupils	9,474	9,474	9,536	9,536

*p<0.1; **p<0.05; ***p<0.01.

All standard errors in brackets, clustered at the school level and corrected for hetereoscedasticity.

Missing dummies have been included for routine teaching practices, pupil/teacher/school characteristics in the regressions.