Forthcoming in the NCAER-Brookings *India Policy Forum 2013*

Priorities for Primary Education Policy in India's 12th Five-year Plan

Karthik Muralidharan[1]

4 April, 2013

Abstract: India has demonstrated considerable progress in the past decade on improving primary school access, infrastructure, pupil-teacher ratios, teacher salaries, and student enrollment. Nevertheless, student learning levels and trajectories are disturbingly low. The past decade has also seen a number of high-quality empirical studies on the causes and correlates of better learning outcomes based on large samples of data and careful attention paid to identification of causal relationships. The findings from this research are however, not being reflected in the current policy priorities of the Government of India. This paper seeks to bridge the gap by summarizing the research, making policy recommendations based on this research, and suggesting an implementation roadmap for the 12th Plan. The main findings reported in this paper are that there is very little evidence to support the notion that improving school inputs in a 'business as usual' manner will improve learning outcomes. On the other hand, innovations in pedagogy (especially supplemental remedial instruction targeted to the level of learning of children) and governance (focused on teacher performance measurement and management) have shown large positive impacts on student learning. The research over the past decade suggests that increasing inputs to primary education in a 'business as usual' way are unlikely to improve student learning in a meaningful way unless accompanied by significant changes in pedagogy and/or improvements in school governance.

JEL Codes: H41, H77, I21, I22, I25, I28, J41, M52

Keywords: Primary Education, India, Research Summary, Policy Recommendations

[1] Department of Economics, University of California San Diego, NBER, and J-PAL; E-mail: kamurali@ucsd.edu.

I am grateful to Abhijit Banerjee, Rukmini Banerji, Jean Dreze, and Parth Shah for their written comments on an earlier draft, and to participants at the India Policy Forum for their comments during the discussion following the presentation of the paper. I thank Shekhar Shah for the invitation to write this paper, which grew out of a lecture at NCAER in December 2011 on "Achieving Quality Universal Primary Education in India: Challenges & Opportunities." I especially thank Montek Ahluwalia, Narendra Jadhav, and Pawan Agarwal for the incorporation of several of the recommendations in this paper into the final version of the Education Chapter of the 12th Five-Year Plan. This paper is based on primary research over the past decade that was enabled and supported by several individuals and organizations. In particular, I thank the Government of Andhra Pradesh, the Azim Premji Foundation, the World Bank, and DFID for enabling the several papers made possible by the "Andhra Pradesh Randomized Evaluation Studies (APRESt)" project. I also thank Pratham, J-PAL, and the state governments that have supported their programs and evaluations for generating several of the insights that inform this paper. I thank Abhijit Banerjee, Rukmini Banerji, Jishnu Das, Shanta Devarajan, Esther Duflo, Eric Hanushek, Caroline Hoxby, Devesh Kapur, Asim Khwaja, Geeta Kingdon, Michael Kremer, Pratap Mehta, Gulzar Natarajan, Lant Pritchett, Halsey Rogers, I V Subbarao, and Venkatesh Sundararaman for the comments, conversations, and collaborations that have informed my thinking on this subject over the past decade. All views expressed are my own.

1. Introduction:

Investing in education is arguably one of the most critical components of enabling the "Inclusive Growth" agenda of the Government of India. Among the several studies carried out on the correlates of long-term economic growth in the nineties, the correlation between average years of education in a country and its growth rate has been among the most robust (Barro 1991, and Benhabib and Spiegel 1994 provide evidence in a cross-country growth regression framework; Mankiw, Romer, and Weil 1992 do so in a growth accounting framework). Concurrently, micro-evidence on the returns to education consistently finds positive returns to primary education in developing countries ranging from 7 to 10% per extra year of schooling (Duflo 2001; Duraisamy 2002). Thus, investments in education are essential for aggregate economic growth as well as for enabling citizens to participate in the growth process through improved wages and employment.

At the same time, recent evidence suggests at both the macro and micro levels that what matters for both growth as well as employability are not years of education as much as the quality of education represented by learning outcomes and skills. In an influential set of papers, Hanushek & Woessmann (2008, 2010) show that cognitive skills as opposed to years of schooling are more robustly correlated with economic growth. They show that the share of basic literates as well as the share of high performers has independent and significant effects on growth and that these types of human capital complement each other. While the results above are based on cross-country regressions, Schoellman (2012) presents micro-evidence using wages of immigrants to the US and shows that cross-country differences in education quality are as important as cross-country differences in years of schooling in accounting for differences across countries in output per worker.

In addition to being an engine of productivity and growth, education quality also determines the extent to which citizens can broadly participate in the growth process. It is a common refrain among employers in India that the majority of college graduates are not 'employable' due to a lack of skills commensurate with their paper qualifications. The weak correlation between years of education and actual knowledge is even more pronounced at the primary schooling level (see section 2). However, while India has made considerable progress in improving primary education when measured by the quality of schooling inputs (including

student enrollment and retention), the progress on learning outcomes has been minimal. It is therefore an urgent priority for primary education policy in India to improve the quality of education measured not just in terms of inputs and student enrollment/retention, but also in terms of learning outcomes.

The past decade has also seen a growing body of high-quality empirical research on primary education in India that can inform primary education policy in a meaningful way. However, the current policy framework for primary education in India (including those in the Right to Education Act) does not reflect the insights from this body of research. The main purpose of this paper is to bridge this gap by distilling the insights from rigorous academic research based on large samples and careful attention to identifying causal relationships, and pointing out the policy priorities that the evidence points towards. This paper does not seek to conduct a comprehensive academic review of this literature with a detailed discussion of econometric identification issues. Rather, it seeks to present education policy makers in India at both the Centre and State-level with a succinct summary of the most credible quantitative research on education over the past decade and then focus on drawing out and discussing the policy priorities suggested by the evidence.[2] In the interests of keeping the scope of this paper manageable, one area that will not be covered is private schools and the optimal structure for leveraging and regulating non-state actors in primary education.[3]

The paper is organized into 4 main sections. Section 2 provides a concise statement of the main facts regarding primary education in India; Section 3 reviews the evidence on the impact of various sets of education inputs (at the school, teacher, and student level) on learning outcomes, reviews the evidence on attempts to improve outcomes by reforming pedagogy and school governance, and finally briefly reviews the evidence on demand-side interventions; Section 4 outlines the policy priorities and approaches for primary education in the coming

[2] The policy recommendations made in this paper reflect the author's judgment of the appropriate weight to be placed on various sources of evidence over the past decade as well as extensive field experience during primary education research in India over this period. For another recent policy paper that summarizes the recent evidence, see Mukerji and Walton (2012), who address similar issues with a more explicit focus on the RtE Act.

[3] The author has ongoing research in the field based on a large multi-year randomized experiment on the causal impact of private schools in India on learning outcomes, and would like to defer the discussion on private schools till we have better evidence. Suggestive evidence on private schools in India based on cross-sectional data is provided in Muralidharan and Kremer (2008) and Desai et al. (2009).

decade suggested by the evidence. Section 5 provides a discussion of implementation challenges and feasible strategies for overcoming these, followed by a brief conclusion.

2. Facts on Primary Education in India

2.1. School quality as measured by inputs has improved considerably in the last decade

A positive consequence of the substantial attention paid to primary education under the past decade by the government of India as well as state governments under campaigns such as the Sarva Shikhsa Abhiyan (SSA) has been the considerable improvement in the quality of government schools as measured by the availability of various kinds of inputs. This can be seen in the trends in the DISE data between 2004 and 2010.[4] In addition to seeing changes in school facilities and teacher quality and quantity in official government reported data, these improvements are also confirmed in data collected completely independent of the government.

Muralidharan, Das, Holla, Kremer, and Mohpal (2013) present results from an all-India panel study of village schools that revisited the rural sample of the nationally representative school survey conducted in 2003 as part of the nationwide study on teacher absence reported in Kremer, Muralidharan, Chaudhury, Hammer, and Rogers (2005). Muralidharan et al (2013) report very significant improvements in input-based measures of schooling quality from this nationally representative panel data. For instance, pupil-teacher ratios have fallen by nearly 20% (from 47.4 to 39.8); the fraction of schools with toilets and electricity has more than doubled (from 40% to 84% for toilets and 20% to 45% for electricity); the fraction of schools with functioning midday meal programs has nearly quadrupled (from 21% to 79%); and the overall index of school infrastructure has improved by 0.9 standard deviations (relative to the distribution of the school infrastructure index in 2003). At the same time, school enrollment rates have increased steadily to the point that 96.7% of children aged 6-14 are now enrolled in school (Pratham 2012).

These are considerable achievements, and should not be regarded lightly given the scale of the Indian primary education system, which is the largest in the world. It highlights that the

[4] Indeed, the investments in high quality administrative data on schools and the creation of the Education Management Information Systems (EMIS) under which the DISE data are made available has also been a significant positive feature in education administration in the past decade.

Indian state does have capacity to execute on goals when undertaken in a "mission mode". These results also suggest ground for optimism that the Indian state is able to make progress on outcomes that are measured and made into a policy priority. However, as we will see below these improvements in school quality as measured by inputs have not translated into improvements in learning outcomes, which may be partly explained by the fact that education policy in the past decade has not prioritized learning outcomes.

2.2. Student learning levels are disturbingly low

While the most prominent set of public discourses on the state of Indian primary education (including those leading up to the Right to Education law) have focused on the low quality of school inputs and schooling conditions (most notable among these was the PROBE Report published in 1999), a new wave of discourse focused on the levels of learning was initiated by Pratham with the publication of the Annual Status of Education Report (ASER) in 2005. This has now become an annual exercise that measures learning outcomes of school-age children in nationally representative samples, with samples large enough to estimate learning levels precisely at the district level.

However, unlike measures of school quality based on inputs (which have shown an upward trend) the picture here is bleak. The most recent ASER report (Pratham 2012) finds that less than 50% of children who are enrolled in the fifth standard are able to read a simple paragraph at the second-standard level, and that less than 27% of children enrolled in the third standard are able to solve a two digit subtraction problem with borrowing and less than 55% of children enrolled in the fifth standard are able to solve the same problem. Over the years, the ASER data suggest that not only are the levels of learning low, but that the trends in learning levels are in fact negative. Since basic reading and arithmetic are foundational skills, the low levels of learning suggested by the ASER data are especially alarming since they suggest that the Indian education system is doing well at enrolling children in school, but failing when it comes to teaching them even basic skills (Pratham 2012).[5]

[5] These figures are based on representative household surveys, and present average achievement levels regardless of whether a student attends a private or a government school. When the figures are broken down by school type, the data consistently show that students in private schools score higher on every measure. Thus, the learning levels for students in government schools are even lower than the ones reported above.

The ASER testing tools are meant to enable a rapid assessment of learning levels and do not span the full range of question difficulty representing the syllabus. It is useful therefore to also look at results from the nationwide School Learning Study conducted in 2010 (Educational Initiatives 2010) by Educational Initiatives, who are one of India's leading testing and assessment firms. These assessments included a broad range of questions including publicly released items from the international TIMSS tests, which would enable a global comparison. The main findings here are consistent with those from the ASER reports. Learning levels are low, and in particular scores on questions that require application of concepts are consistently lower than those on questions representing rote learning. The report also finds that the mean score across Indian public schools on the common TIMSS questions in the standard 4 language test is less than half that of the international mean (less than 30% compared to over 60%).[6]

Muralidharan and Zieleniak (2013) use a unique longitudinal data set in the state of Andhra Pradesh collected by following a cohort of students over five years and find that not only are learning levels low, but so are the learning trajectories over time. They use item response theory (IRT) to create item characteristics of a 3-parameter logistic model (difficulty, discrimination, and guessing parameters) for a database of over 900 questions each in math and language that were administered as part of the APREStudies over five years. Using overlapping questions over years and a set of identical questions that were administered simultaneously to students across grades 1 to 5, they estimate learning trajectories, defined as the probability of a typical student in a given grade getting a question correct over time as they progress through the grades. Their findings suggest that for most questions of intermediate levels of difficulty, less than 20% of students who do not correctly answer a grade N-level question at the end of grade N, are able to answer it correctly at the end of grade N+1. These results suggest that spending additional years in school, while no doubt useful in terms of added learning, has remarkably low effectiveness in improving learning outcomes - especially given the considerable economic cost of an additional year in school. They also find evidence of increasing variance in absolute learning levels of students over time.[7]

[6] The results are not reported in standard deviations (Educational Initiatives 2010).
[7] Note that this probably understates the increase in variance because of a higher probability of students dropping out from the lower end of the learning distribution.

The studies mentioned above are all unanimous in suggesting that learning levels in India are low by any absolute standard. But the magnitude of India's 'learning deficit' is particularly stark when placed in an international comparative context. Das and Zajonc (2010) show that learning levels in the Indian states of Orissa and Rajasthan would fall below 43 of the 51 countries for which comparable TIMSS data are available. Even more striking is the finding of the recent PISA assessments carried out in two of the more advanced Indian states in terms of learning levels – Himachal Pradesh and Tamilnadu – which finds that the 2 tested Indian states ranked 72nd and 73rd out of a total of 74 tested entities for which results were reported (not all were countries). Combining these results with those of the SLS (2010) suggests that many of the more educationally backward states like UP, Bihar, and Jharkhand would lag even further behind in international comparisons (and drag down the population-weighted all-India means much further). It is worth highlighting that these results do not simply reflect the correlation between economic development and test scores because the top scoring entity was the city of Shanghai in China, which has the annual per-capita income of a middle-income country (approximately 13,000 US Dollars per head as of 2011, which is comparable to that of Brazil).

Thus, while the quality of schooling as defined by traditional notions of school inputs has been improving steadily due to increased government expenditure, quality as defined by learning outcomes is low both in absolute terms (measured by what competencies children in school are demonstrating) as well as in relative terms (as seen in the PISA scores).

2.3. There is an increasing and widespread exodus to fee-charging private schools

There is perhaps no greater indicator of the quality of government schooling as perceived by parents than the increasing extent to which parents are eschewing free government schools (in fact government schools have a 'negative' cost once the various incentives such as mid-day meals, free text books, and other benefits are accounted for) and moving their children to fee-charging private schools. Desai, Dubey, Vanneman, and Banerji (2009) show, using nationally-representative data from 2005, that 58% of students in urban India attended fee-charging private schools. The annual ASER reports show a steadily increasing trend in private school enrollment from 18.7% in 2006 to 25.6% in 2011 – with these increases being broad based across states. These numbers highlight that India has a share of private school enrollment that is comparable to a country like Chile – that has a fully voucher-based school system!

It is beyond the scope of this paper to compare the effectiveness of private and government-run schools, but these data indicate that in spite of considerable increases in spending on government schools, parents do not perceive this spending to be generating enough quality in the government schooling experience for them to retain their children there. While it is true that parents value many things in schools (with learning outcomes being only one component in a vector of schooling attributes that parents care about), the trend towards increasing private school share in primary education combined with the low levels of learning outlined in the previous section suggest that there are considerable systemic weaknesses in translating increasing education spending into superior outcomes in government-run schools.

3. Reviewing the Evidence on Causes and Correlates of Learning Outcomes

The main factors that determine the performance of a school system include the level of inputs provided (facilities, teachers, and student inputs), the pedagogical processes employed in classrooms, and the overall governance of the school system. In addition to these supply-side factors, a further key determinant of educational attainment is the extent of demand for education from parents and students. Each of these areas has seen considerable empirical research in the past decade and this section briefly summarizes the evidence on these broad classes of issues that are relevant to the translation of spending into outcomes.

3.1. Inputs

The most important components of education spending in the past decade have been on improving school facilities and infrastructure, improving teacher salaries and training, hiring more teachers to reduce pupil-teacher ratios, and expenditure on student benefits such as textbooks, and mid-day meals. The PAISA Report (Accountability Initiative 2012) shows that these three categories of expenditure account for 90% of the SSA budget (in the most recent year, 44% was spent on teachers, 36% on schools, and 10% on students - though the last category does not include spending on mid-day meals). However, as the discussion below shows, the empirical studies to date do not find significant correlations between these investments and either intermediate measures of system performance (such as teacher absence) or measures of outcomes (such as student test scores).

3.1.1. School Infrastructure

In the absence of rigorous randomized evaluations studying the impact of infrastructure improvement on learning outcomes in India, the broadest evidence to date comes from Muralidharan et al (2013). Using village-level panel data from a nationally-representative sample of over 1,250 villages across 19 Indian states, they find no correlation between changes in average village-level school infrastructure (between 2003 and 2010) and changes in enrollment in government schools, though they do find a small positive effect on the number of students attending school. They also find no correlation between changes in average village-level school infrastructure and either teacher absence or student test scores, even though as noted earlier they find significant improvements in almost all measures of school infrastructure.

One experimental evaluation of an infrastructure intervention is Borkum, He, and Linden (2010) who study the impact of a school-library program in Karnataka. They find that even though the program provided schools with several new books as well as a librarian, the program had no impact on student reading scores. Analysis using the five-year panel data set of student learning outcomes collected as part of the APREStt project also finds no correlation between the infrastructure index in the school and measures of student test-score gains.[8]

Thus, almost all the existing evidence points to a limited impact of improvements in school infrastructure on learning outcomes. The reasons for this are not obvious. One possibility is that these investments make schools more appealing to teachers and students, but have no impact on the teaching and learning process - which may be the main determinant of learning. Another possibility is that infrastructure may be built but not used. For instance, the APREStt project collected matched data between school facilities and household behaviors and the data suggests that over 75% of children who attend schools that have a toilet still report relieving themselves in the open in school.[9] A final possibility is that the returns to infrastructure investments need to be evaluated over the depreciation lifecycle of the corresponding infrastructure. It is possible that the cumulative impact of investments in buildings over a 30-

[8] Calculations by author using the APREStt data. Note that these are not experimental results, but by controlling for lagged test scores, this analysis mitigates several of the usual omitted variable concerns.
[9] This could be for logistical reasons such as lack of water in the school toilet or the lack of staff to clean the toilet, due to which teachers may prefer to keep the toilets closed. Alternatively, these results could reflect the difficulty of changing behavioral norms with respect to sanitation.

year depreciation lifecycle may be significantly positive, while the annual effect on learning outcomes is too small to be measured statistically.

This last possibility should caution us against interpreting the results to date as suggesting that infrastructure investments should not be made. More broadly, the results should not be interpreted as saying that school infrastructure does *not* matter for improving learning outcomes (they may be necessary but not sufficient), but the evidence does suggest that investment in infrastructure by *itself* is unlikely to have a significant impact on improving learning levels and trajectories. This is essential to point out because the staffing patterns of education department offices around the country suggest that the dominant concern for the department is typically infrastructure and facilities, while there are almost no staff at the district and block levels whose main task is to focus on academics and pedagogy.[10]

3.1.2. Teacher Quantity and Quality

The other major component of investment in inputs has been increasing teacher salaries and training, and reducing pupil-teacher ratios. The evidence summarized below again points to very limited impacts of these investments on improved learning outcomes.

While there has been no experimental evaluation of the impact of varying individual teacher characteristics in India, there have been quite a few studies that control for lagged test scores and estimate the impact of teacher characteristics on learning outcomes in a value-added framework. The first point to highlight is that *none of these studies* to date finds a significant positive relationship between teacher training and increases in test scores of students taught by the corresponding teacher (see Kingdon and Teal 2010; Muralidharan and Sundararaman 2011b, 2013; Muralidharan 2012). Similarly, there is no correlation between teacher salary and student test score gains (Kingdon and Teal 2010; Muralidharan and Sundararaman 2011b; Muralidharan 2012)[11], and if anything the correlations typically point to a *negative* relationship between teacher salaries and gains in student test scores.

[10] Thanks to Rukmini Banerjee for highlighting this point in her discussion.
[11] The results from Muralidharan and Sundararaman (2011b) and Muralidharan (2012) referred to here are based on the tables of heterogeneous treatment effects of the performance-pay interventions as a function of teacher characteristics. The specifications used our standard value added specifications and the results reported above are

The evidence on the impact of reducing pupil-teacher ratios on improved learning outcomes is also quite mixed, with most studies not finding much of an impact. Banerjee, Cole, Duflo, and Linden (2007) report results from an experimental evaluation that provided remedial instruction to children with low test scores by taking them outside the regular classroom for remedial instruction provided by a volunteer. However, while the test scores of the children who received this remedial instruction went up significantly, they find no impact on the test scores of the students who remained in the original classroom with a smaller class size. These results suggest that reducing class-size may have a limited impact on improving test scores.

Muralidharan and Sundararaman (2013) study the impact of school level pupil-teacher ratio (PTR) on test score gains by using longitudinal data on test scores and changes in PTR over time and find significant but modest gains from reducing the school level PTR. Their estimates imply that reducing school level pupil-teacher ratio by half would at most yield gains in test scores of 0.25 standard deviations per year. Jacob, Kochar, and Reddy (2008) study the impacts of class size on learning outcomes on Andhra Pradesh using a control-function approach and also find significant but small effects of class-size reductions on test scores.

Further, the panel data analysis conducted by Muralidharan et al (2013) finds no correlation between changes in mean pupil-teacher ratio in a village and changes in normalized mathematics test scores. They also find evidence of a possible mechanism for this finding, which is that there is a very robust *negative* relationship between pupil-teacher ratio and teacher absence. In other words, *reductions* in pupil-teacher ratio over time were strongly correlated with *increases* in teacher absence. Thus, the impact of reducing class size by hiring additional teachers was mitigated by increased levels of teacher absence in the schools. This is consistent with the experimental evidence presented in Muralidharan and Sundararaman (2013) where they find that schools that were randomly selected to receive an additional contract teacher saw a significant increase in the absence rates of the regular teachers.[12] In other words, the marginal rate of teacher absence may be considerably higher than the average, which could limit the impact of reducing pupil-teacher ratio on improving learning outcomes.

---

the coefficient on the linear term (the main effect of the characteristic) and not the interaction term (which measures the heterogenous impact of the performance pay program as a function of the characteristic).

[12] Similar findings are reported by Duflo, Dupas, and Kremer (2012) in an experimental study of contract teachers in Kenya, suggesting that this may be quite a general result.

Finally, a related issue is the one of distribution of teachers across schools. While budgetary considerations lead to a focus on average PTR's, in practice there is wide variation in PTR's across schools. Chin (2005) shows that Operation Blackboard in India which redistributed teachers from large to small schools led to a significant increase in primary school completion rates for girls and the poor even though there was no increase in the average number of teachers per school and no reduction in mean class size.

Summarizing the research on PTR on learning outcomes, we see that the best studies do find some positive impacts of class-size reduction on student test scores. Nevertheless, these estimated impacts are modest in magnitude, and given the high cost of class-size reductions, it may not be very cost effective to aim to improve test scores by reducing class sizes. Thus even a 20% reduction in pupil-teacher ratio (which is a very expensive intervention) would not yield large test score gains (around 0.05 standard deviations/year) and would be considerably less cost effective than achieving the same class-size reduction using contract teachers (Muralidharan and Sundararaman 2013) or introducing modest amounts of performance linked bonuses (Muralidharan 2012; see section 3.3.4). The evidence also suggests that in addition to average PTR's, it may also be important to pay attention to the distribution of teacher resources across and within schools, and that it may be possible to improve learning outcomes at no additional cost simply by rationalizing the allocation of teachers across schools, and by providing smaller class sizes to earlier grades.

3.1.3 Student Grants and Mid-Day Meals

The final major category of inputs is student-based spending including textbooks, uniforms, and mid-day meals. Again, studies to date do not find any significant positive relationship between these categories of spending and improved learning outcomes.

Das, Dercon, Habyarimana, Krishnan, Muralidharan, and Sundararaman (2013) present experimental evidence on the impact of a school grant program that stipulated that the funds should be spent on inputs directly used by students. The program was implemented over two years in the major categories of spending were books, stationery, and writing materials (~50%); workbooks and practice books (~20%); and classroom materials (~25%) with similar patterns of expenditure in both years of the program. They find that this program had a significant positive

impact on student test scores at the end of the first year, but that the impact in the second year was close to zero, with the cumulative two-year effect being positive but not significant. They show the most likely mechanisms of this result is that households considerably reduce their own spending on their child's education in the second year of the program.

Thus, when the program was unanticipated and when the money arrived after parents had already incurred their educational expenditures on books and materials for the school year (as in the first of the program), there was a significant net increase in materials which translated into significant improvements in test scores. However, when these inputs were anticipated, households were able to re-optimize and reduce their own spending. Thus, there was no significant increase in net inputs in the second year which would explain why there was no impact on test scores either. These results highlight the importance of accounting for household re-optimization in response to public spending programs in thinking about the long-term impacts of increased spending, and suggest a possible mechanism for the lack of correlation between increased spending on inputs and improved outcomes.[13]

A similar concern exists in the context of mid-day meals, because it is possible for households to adjust the allocation of food within the household in response to the fact that the school going child now has access to one meal in the school. Afridi (2010a) studies the impact of mid-day meal provision and finds that the program substantially increases the total caloric intake of school-going children in rural Madhya Pradesh, by 50% to 100%. Using a difference-in-difference estimation strategy that relies on a staggered roll out across schools, attendance rates for girls are estimated to increase by 12 percentage points in rural Madhya Pradesh (Afridi, 2010b) and 5 percentage points overall in Delhi (Afridi, Barooah and Somanathan, 2010). However, these papers do not study the impact of mid-day meals on test scores. Jayaraman, Simroth and Vericourt (2010) use data from thirteen states to construct triple-difference estimates using private schools as a control group and find that the mid-day meal program is associated with a 6.8% increase in enrollment, but had no impact on test scores. Finally, the panel data analysis in Muralidharan et al (2013) finds that there is a *negative* (though not always

[13] In technical terms, these results highlight that it is possible for the production function effect of additional inputs on test scores to be positive (this is a partial derivative of the impact of additional inputs holding other factors constant), while the policy effect might be considerably lower (since this includes re-optimization by other agents). This is clearly a very general theme since the discussion in the previous section of increased absence among pre-existing teachers in response to the addition of a new teacher is an illustration of the same point.

significant) correlation between changes in the mid-day meal status of schools in a village, and changes in normalized math test scores. One possible mechanism for this result may be the diversion of teacher time to manage and oversee the mid-day meal process. Analysis of teacher time use data in Andhra Pradesh using the APRESt data, suggests that government school teachers report spending around 10% of their daily time in school overseeing the mid-day meal.

Another student input that has been found to have a significant impact on enrollment, but insignificant impact on learning outcomes is the bicycles that have been provided to girls in several states to improve secondary school enrollment. Muralidharan and Prakash (2013) study the impact of the Chief Minister's Bicycle Program that provided girls in Bihar with a bicycle conditional on enrolling in 9th grade. They use a triple difference approach (using boys and the neighboring state of Jharkhand as comparison groups) and find that the program increased girls' enrollment in secondary school by 20% (a five percentage point gain on a base enrollment rate of twenty-five percent) and reduced the gender gap in secondary school enrollment by 25%. They find that the impact of the program was significantly greater in villages where the nearest secondary school was further away, suggesting that a key mechanism for program impact was the reduction in the 'distance cost' of school attendance induced by the bicycle. However, they do not find any significant impact of the cycle program on girls' learning outcomes as measured by their passing rates in the tenth-standard board exam.

To summarize, it appears that most of the investments in improving school quality as measured by inputs (regardless of whether these are at the school, teacher, or student level) are either not correlated with improved learning outcomes or only weakly so. There may well be other important reasons for making these investments (such as child welfare), and student inputs that reduce the marginal cost (or increase the marginal benefit) of attendance do seem to have a positive impact on school participation. But the evidence to date does not suggest any reason to be optimistic that 'improving' school quality in a 'business as usual' way will lead to a substantial improvement in learning outcomes.

3.2. Pedagogy

While there have been significant increases in schooling inputs, a key determinant of how these investments translate into learning outcomes is the structure of pedagogy and classroom

instruction. Getting aspects of instruction right is particularly challenging in a context such as India where several millions of first-generation learners have joined a rapidly expanding national schooling system. In particular, standard curricula and teaching practices that may have been optimal at a time when education was more limited may not fare as well under the new circumstances. The discussions in this section focus on some key aspects of classroom structure and pedagogy that are relevant for the South Asian context – including remedial instruction, and the use of technology in the classroom.

3.2.1. Remedial Instruction

A fundamental challenge for pedagogy in a context of several millions of first-generation learners is the large variation this creates in the initial preparation of children when they enter school. Also, as Muralidharan and Zieleniak (2013) show, the variance in student learning levels increases over time. How does a teacher effectively teach a classroom where students are so varied in their skill level? Remedial schooling interventions have been one method to attempt to reduce the variance of achievement in the classroom and ensure that all students are progressing. Remedial programs offer the possibility of focusing on those students who are lagging behind and teaching at a level that is appropriate for their achievement. Ideally, such an intervention would increase their progress, and decrease the heterogeneity of student learning levels in a given grade.

The evidence confirms that this may be the case, with several high-quality studies finding strong impacts of remedial instruction programs on learning outcomes, even when implemented by volunteers or informal teachers with little formal training and paid only a modest stipend that is several times lower than the salary of regular government teachers.

First, Banerjee, Cole, Duflo, and Linden (2007) report results from an experimental evaluation of a program run by PRATHAM specifically targeted at the lowest performing children in public schools in the Indian cities of Mumbai and Vadodara. The program provided an informal teacher hired from the community (known as a Balsakhi or 'friend of the child') to schools, with an explicit mandate to focus on children in 3rd and 4th grade who had not achieved even basic competencies in reading and arithmetic. These children were taken out of the regular classroom for 2 hours a day, and were provided with remedial instruction targeted at their current

level of learning. The program improved student test scores by 0.28 standard deviations, with most of the gains coming from students at the lower end of the learning distribution.

Second, Banerjee, Banerji, Duflo, Glennerster, and Khemani (2010) report results from several interventions designed to improve community participation in education. Of all the interventions tried, the only one that was found to be effective at improving learning outcomes was a remedial instruction program implemented by youth volunteers hired from the village who were provided a week of training and conducted after school reading camps for two to three months. These effects were substantial (albeit off a low base) with the average child who was not able to read anything at the baseline and who attended a camp being 60 percentage points more likely to be able to read alphabets than a similar child in a control village.

A third piece of experimental evidence is provided by Lakshminarayana, Eble, Bhakta, Frost, Boone, Elbourne, and Mann (2012), who study the impact of a program run by the Naandi Foundation that provided remedial education program run by community volunteers to a randomly selected set of villages in Andhra Pradesh. After an initial sensitization to households regarding the program, the volunteers provided two hours a day of remedial instruction after normal school hours in the school itself (on a daily basis). The subject matter covered in these sessions was tailored to students' class-specific needs and learning levels, and aimed to reinforce the curriculum covered in school. At the end of two years of this intervention, student test scores in program villages were 0.74 standard deviations higher than those in the comparison group, suggesting a large impact of the after-school remedial instruction program.

Finally, Banerjee, Banerji, Duflo, and Walton (2012) study the impact of a program implemented by Pratham in partnership with the state governments of Uttarakhand and Bihar that attempted to scale up remedial instruction in public schools, and find that summer camps conducted by regular teachers transacting the learning-appropriate remedial materials were effective in raising test scores. However, they find that there was no impact of other models that attempted to incorporate this pedagogy in the regular school day. The authors interpret their findings as suggesting that the remedial pedagogy was successful, but that it was difficult to get teachers to implement new curriculums during school hours.

3.2.2. Technology-Aided Instruction

Greater use of technology in classrooms is commonly thought of a promising way to rapidly improve education outcomes in developing countries (including India). Posited channels of impact include (1) cost-effective replication and scaling up of high-quality instruction using broadcast technology (such as radio and television-based instruction); (2) using technology to overcome limitations in teacher knowledge and training (for instance for teaching more advanced concepts in science and mathematics or for teaching a new language like English – for which there is growing demand but a limited supply of teachers with the requisite competence); (3) using technology to provide supplemental instruction at home; (4) using technology to engage children better in the learning process through the use of interactive modules (such as educational games and puzzles); and (5) using technology to customize individual student learning plans. These interventions also range from being quite inexpensive on one hand (radio-based instruction for instance) to very expensive (individual laptops for students such as envisaged under the 'One Laptop per Child' or OLPC initiative).

While the promise of enhanced use of technology in instruction is clear, and there are many advocates for doing so, the evidence on the effectiveness of technology in instruction remains limited and few rigorous studies have evaluated the benefits of such interventions. Skeptical scholars have even argued that the promotion of technology is fueled more by the prestige and symbol of modernity than any actual evidence of the effectiveness of the interventions (Shields 2011). While many continue to champion educational technology, there may be adverse consequences of their implementation, the simplest of which would be an ineffective technology that does not increase achievement and takes time away from other more effective teaching techniques. Understanding the efficacy of technology is especially important as technology is often relatively expensive compared to other activities – if they do not lead to superior learning outcomes, then it is likely that there are more cost effective methods than technology to improve educational outcomes.

Linden (2008) evaluates the impact of a computer-aided instruction program implemented by an NGO in Gujarat (Gyanshala) that was implemented both in an after-school supplemental instruction model as well as in a model where computer-aided instruction replaced a period of regular instruction. The paper finds that the supplemental program led to significant

positive effects on test scores (0.28 standard deviations), while the in-school model led to significantly lower test scores (-.0.57 standard deviations) suggesting that a blanket use of 'computers in school' may not only not be effective, but could also be harmful if it replaces otherwise productive instructional time.[14]

Further evidence on the importance of design details is provided by He, Linden, and MacLeod (2008) who analyze an intervention aimed at improving English skills in which part of the intervention is directed by teachers and the other component is a self-paced machine. While both components led to positive gains in test scores, the study found that stronger students fared better using the machine, while weaker students benefited more from the guidance of a teacher. Thus, technology may be an effective teaching aid, but it may require higher initial levels of learning to be used effectively.

Banerjee et al (2007) find that a computer remedial program increases test scores twice as much as the remedial teacher. However, because of the high expense of the computer-based program, scaling up the teacher-based remedial program would be 5 to 7 times more cost effective than the computer assisted learning program. The experiment illustrates that while certain technologies may be effective, it still may be more cost effective to use non-technology based programs.

Finally, while set in a different middle-income context, it is worth highlighting results from an experimental evaluation of the much-publicized "One Laptop Per Child (OLPC)" program in Peru (Cristia, Ibarraran, Cueto, Santiago, and Severin 2012). The paper finds that while the program increased the ratio of computers to students in schools from 0.12 to 1.18 in treatment schools, there was no impact on either school enrollment or test scores in Math and Language. The paper does find some positive effects on general purpose measures of intelligence such as the Raven's Progressive Matrices but the overall results suggest need for caution in believing that the introduction of computers in classrooms will by itself lead to improvements in learning levels.

[14] While set in a different context, a well identified study on the impact of providing 14-year old students with computers at home in Romania also found negative effects of the computer on test scores (Malamud and Pop-Eleches 2011) – again serving to caution that a naïve attempt to provide students with more technology can have negative effects and that interventions need to pay careful attention to what activities are being crowded out by the additional computer time.

These cautionary results are especially relevant in a context such as India where it is tempting to scale up interventions like "tablet computers for all" as a potential short-cut for addressing the challenges of education quality. To summarize, there is are many good reasons to be excited about the *potential* for technology-enabled instruction to improve learning outcomes significantly. However, the evidence on the impact of greater use of technology in the classroom is mixed and seems to depend crucially on the details of the model by which it is implemented. A lot more careful research is needed (on both process and impacts) before committing resources to scaling up these programs - especially those involving expensive investments in hardware.

3.3. Governance

Beyond pedagogy, another explanation for the low correlation between increases in spending on educational inputs and improved learning outcomes may be the weak governance of the education system and limited effort on the part of teachers and administrators to improve learning levels. This section reviews the evidence on some of the key themes relating to school governance in India.

3.3.1. Teacher Absence

Perhaps the most striking measure of weakness of school and teacher governance in India is the high rate of teacher absence from schools. Kremer et al(2005) present results from a nationally-representative all-India survey of schools where enumerators made unannounced visits to schools to measure teacher attendance and activity. They find that on any given day, around 25% of teachers were absent from work, and less than half of the teachers on the payroll were found to be engaging in teaching activity. The absence rate was the second highest in a similar survey across 8 low and middle income countries.

Muralidharan et al (2013) present results from a nationally-representative panel survey that revisited the villages visited in the study above, and find that there has been a reduction in teacher absence rates from 26.3% to 23.7%.[15] While this is a significant reduction in teacher absence rates, the magnitude of improvement in measures of governance such teacher absence is considerably lower (0.26 standard deviations relative to the 2003 distribution of teacher absence)

[15] The absence rate of 25% includes both the rural and the urban sample, whereas the absence rate in the rural sample in 2003 was 26.3% (for the villages in the panel data set)

than the magnitude of improvement in physical inputs such as school infrastructure (0.91 standard deviations relative to the 2003 distribution).

In addition to these 2 nationally-representative studies, several other studies have also noted the high rates of teacher absence in India. Duflo, Hanna, and Ryan (2012) find teacher absence rates in excess of 40% in informal schools run by an NGO in Rajasthan. Muralidharan and Sundararaman (2011b, and 2013) , and Muralidharan (2012) regularly document teacher absence with multiple unobserved visits to a representative sample of rural government-run primary schools in Andhra Pradesh and find teacher absence rates to steadily range between 24 - 28% over the 5 year period from 2005-06 to 2009-10.

3.3.2. Monitoring

Muralidharan et al (2013) use their nationally-representative panel data set on teacher absence to estimate the correlations between changes in various school and management characteristics from 2003 to 2010 and changes in teacher absence. Among all the variables they study, there are only 2 robust correlates of teacher absence that are significant under all specifications (with and without state/district fixed effects). The first is the negative correlation between pupil-teacher ratio and teacher absence (described in section 3.1.2), and the second is the strong negative correlation between school inspections and teacher absence. They find that increasing the probability of a school having been inspected in the past 3 months from 0 to 1 is correlated with a 7 percentage point reduction in teacher absence (or 30% of the observed absence rates). This estimate is similar in both cross-section and panel estimates, bivariate as well as multiple regressions, and with and without state/district fixed effects. Using the most conservative of these estimates, Muralidharan et al (2013) calculate that increasing inspections/monitoring could be over 10 times more cost effective at increasing teacher-student contact time (through reduced teacher absence) than hiring additional regular teachers.

On the other hand, the correlations between 'bottom up' measures of governance and monitoring such as the frequency of PTA meetings and teacher absence is also negative but the magnitude is always lower than that of the 'top down' inspections and is not always significant. These results highlight that there may be significant collective action problems that may make community-based monitoring less effective than top-down administrative monitoring (a result

consistent with the experimental findings of Olken 2007 in the context of monitoring corruption in Indonesia). Banerjee et al (2010) provide experimental evidence on the challenges of using community mobilization to improve school quality. They find no impact of various programs to build community involvement in schools in Uttar Pradesh on community participation, teacher effort, or learning outcomes.

Duflo, Hanna, and Ryan (2012) conduct an experimental evaluation of an intervention that monitored teacher attendance in informal schools in Rajasthan using cameras with time-date stamps to record teacher and student attendance. The program also paid teacher salaries as a function of the number of valid days of attendance. They find that this program reduced teacher absence by half, but structural estimates of a model of labor supply suggest that the mechanism for this result was not the 'monitoring' per se, but rather the incentives tied to the attendance. Muralidharan and Sundararaman (2010) study the impact of a program that provided schools and teachers with low-stakes monitoring and feedback and find that this program had no impact on either teacher attendance or test scores. These results suggest that while 'monitoring' is an important tool in reducing teacher absence, 'low-stakes' monitoring is unlikely to be very effective, and that it is 'high-stakes' monitoring with positive/negative consequences for presence/absence that is more likely to be effective.

3.3.3. Contractual Structure

A widespread but highly controversial aspect of primary education policy in India during the past couple of decades has been the use of locally-hired contract teachers on fixed-term renewable contracts, who are not professionally trained, and who are paid *much lower* salaries than those of regular teachers (often less than one fifth as much).[16] Supporters consider the use of contract teachers to be an efficient way of expanding education access and quality to a large number of first-generation learners, and argue that contract teachers face superior incentives compared to tenured civil-service teachers. Opponents argue that using under-qualified and untrained teachers may staff classrooms but will not produce learning outcomes, and that the use

[16] Contract teacher schemes have been widely employed in several states of India (under different names such as Shiksha Karmi in Madhya Pradesh and Rajasthan, Shiksha Mitra in Uttar Pradesh, Vidya Sahayak in Gujarat and Himachal Pradesh, and Vidya Volunteers in Andhra Pradesh). The salary differentials are even more pronounced if we account for the present discounted value of the pension and other retirement benefits offered to civil-service government teachers.

of contract teachers de-professionalizes teaching, reduces the prestige of the entire profession, and reduces motivation of all teachers.[17] However, as seen below, there is no evidence to support the view that contract teachers are less effective than regular teachers.

Muralidharan and Sundararaman (2013) present experimental evidence from a program that provided an extra contract teacher to 100 randomly-chosen government-run rural primary schools in the Indian state of Andhra Pradesh. At the end of two years, students in schools with an extra contract teacher performed significantly better than those in comparison schools by 0.16 and 0.15 standard deviations, in math and language tests respectively. They also find that contract teachers were significantly less likely to be absent from school than civil-service teachers (16% vs. 27%). Finally, they implement four different non-experimental estimation procedures (using both within and between-school variation as well as variation over time in pupil-teacher ratios in the same school) and find that they can never reject the hypothesis that contract teachers are at least as effective in improving student learning as regular civil-service teachers. In fact, their point estimates typically suggest that the contract teachers are more effective than regular teachers who are more qualified, better trained, and paid five times higher salaries.

Atherton and Kingdon (2010) use data from Uttar Pradesh and estimate the relative effectiveness of contract and regular teachers using a student fixed-effects approach (exploiting variation in the contract/regular teacher status of teachers who are teaching different subjects to the same student) and find that the contract teachers produced better learning outcomes. Finally, Goyal and Pandey (2011) use data from Madhya Pradesh and Uttar Pradesh and find that contract teachers exert higher levels of effort than regular teachers with employment security (on measures of teacher attendance and engagement).

It is also relevant to this discussion to highlight that all the four studies discussed in the previous section that found large positive effects on student learning outcomes of remedial instruction programs, used volunteer/informal/contract teachers with minimal formal training who were paid stipends that were at most one fifth of the salary of regular teachers. These results suggest that the superior work incentives of contract teachers may more than make up for

[17] See Kumar et al (2005) for an example of these criticisms.

their lack of formal teacher training. They also suggest that the binding constraint in translating increased education spending into improved learning outcomes may not be teacher training and qualifications (as is commonly believed) but teacher effort, which is (relatively) weaker for civil-service teachers with lifetime employment security because there is no reward for effort and performance under the status quo (and conversely, few consequences for poor performance).

3.3.4 Performance-Linked Pay

The discussions in this section suggest that improving governance is not just a matter of making better policies but also requires enhancements in the capacity of the government to effectively *implement* policies. Since the effort exerted by public sector employees is a key determinant of state effectiveness, a natural set of policy options to enhance governance in education would be to consider linking compensation of teachers as well as education administrators to measures of performance.

Muralidharan and Sundararaman (2011b) present experimental evidence on the impact of a program in Andhra Pradesh that provided bonus payments to teachers based on the average improvement of their students' test scores in independently administered learning assessments (with a mean bonus of 3% of annual pay). At the end of two years of the program, students in incentive schools performed significantly better than those in control schools by 0.27 and 0.17 standard deviations in math and language tests respectively. Students in incentive schools also performed better on subjects for which there were no incentives, suggesting positive spillovers between improved performance on math and language and the untested subjects (science and social studies). Since the performance pay programs were implemented as a part of larger set of experimental evaluations costing the same amount, the authors are able to compare the relative effectiveness of input and incentive based approaches to improving learning outcomes. They find that the incentive schools performed significantly better than other randomly-chosen schools that received additional schooling inputs of a similar value.

Also, as discussed earlier, Duflo, Hanna, and Ryan (2012) find that paying teachers on the basis of the number of days they attend work (as opposed to a flat salary that does not depend on performance) led to a halving of teacher absence rates (from 42% to 21%) and significant increases in student test scores (by 0.17 standard deviations).

Finally, Muralidharan (2012) presents evidence from the longest-running experimental evaluation of a teacher performance pay program (spanning 5 years), and finds that students who completed their full five years of primary school under the individual teacher incentive program performed significantly better than those in control schools by 0.54 and 0.35 standard deviations in math and language tests respectively. The group teacher incentive program also had positive (and mostly significant) effects on student test scores, but the effect sizes were always smaller than that of the individual incentive program, and were not significant at the end of primary school for the cohort exposed to the program for five years. The paper estimates that the individual teacher performance pay program would be around 15 to 20 times more cost effective (including administrative costs) at improving learning outcomes than the default policy of reducing pupil-teacher ratios by hiring more teachers (even assuming the most generous estimates of the impact of PTR reductions on test scores from the discussion in section 3.1.2).

Taken together, these results suggest that even modest changes to compensation structure to provide reward and recognition to teachers on the basis of objective measures of performance (such as attendance or increases in student test scores) can generate substantial improvements in learning outcomes at a fraction of the cost of a 'business as usual' expansion in education spending.

3.4 Demand Side Interventions

The discussion so far has focused mainly on the supply side of education, since this is what typically concerns what the government does in terms of running schools. However, the amount of education obtained by a child typically reflects a decision made by parents that considers the costs and benefits of education as well as other considerations (including credit, information, discount rates, risk preferences, and time horizon). Indeed it is possible that the sharp increases in school enrollment over the past decade have been driven not so much by the education policies of the government as much as they have been by rapid economic growth and increasing real and perceived *returns* to education, which in turn have boosted the demand for education.[18] Nevertheless, it is possible that there is still under-investment in education because

[18] While there is no research that credibly quantifies the relative importance of supply and demand side factors in improving education attainment in India, there are several studies that highlight the importance of increasing returns

of demand side failures including incorrect perceptions on the returns to education, and high discount rates of parents.

3.4.1. Providing Better Information on Returns to Education

Since household decisions regarding education investments are made on the basis of *perceived* as opposed to actual returns to education, interventions that provide better information about education options and the mean and distribution of outcomes at different levels of education may improve decision making regarding education investments. In a randomized evaluation in the Dominican Republic, Jensen (2010) found that providing eighth-grade boys with information on the returns to secondary education increased the years of education completed by .25-.30 years. In an experimental study in Madagascar, Nguyen (2008) finds similarly large effects on student test scores of simply providing better statistics to students on the mean wages at different levels of education. These gains are remarkable given the simplicity of the intervention, which involved reading a simple statement to students. However, one challenge is that the returns to education are typically not very credibly estimated (especially in countries with rapidly transforming economies – such as India). Also, returns to education are likely to be heterogeneous and accurate estimates of the distributions of returns to education are even more difficult to obtain. These complications raise the risk of providing incorrect information to households regarding returns to education, which may make them worse off.

A good way to address this concern (and still provide useful information) is demonstrated by Jensen (2012) who presents the impact of a program in North India where recruiters for call centers visited villages and hired girls who met the job requirements for working in call centers. He finds that women in treatment villages were significantly less likely to get married or have children during this period, and more likely to either enter the labor market or obtain more schooling. But this intervention provides information on returns to education not by showing average returns calculated from a (potentially incorrect) Mincer regression, but by demonstrating to village residents that girls with a high-school education can get hired by call centers. This is important because the recruiting standards were *not* changed, and so no (potentially) incorrect information was provided. But the intervention did provide accurate new information to village

---

to education in household decision making with respect to educational attainment including Munshi and Rosenzweig (2006), Jensen (2012), and Shastry (2012).

residents regarding the job possibilities for educated girls because the recruiters would typically not have visited the village (since the expected number of recruits would not justify the fixed costs of the recruiters going to the village).

The success of all these information-based interventions suggests that this may be a particularly useful avenue to explore for increasing education participation – especially since information interventions can be carried out relatively inexpensively.[19]

4. Policy Recommendations

While there has been a considerable amount of high-quality research in the past decade on what does and does not seem to matter for improving learning outcomes in India, it is not obvious that each of these individual research findings should directly translate into policy. Policy formulation needs to consider technical, administrative, ethical, as well as political factors and even the best technical studies can only provide inputs into one dimension of policy making. For instance, many programs which may not be 'cost-effective', such as education for children with special needs, may nevertheless be consistent with normative principles of a just and humane society. Nevertheless, given budgetary pressures and the existence of several sectors that can claim an ethical basis for increased spending in a fiscally constrained environment (including health and food security), it becomes both morally and practically imperative to account for cost-effectiveness in questions of public policy. Improving the cost-effectiveness of social sector spending will allow a fiscally constrained state to do more in the social sector and improve both efficiency of spending as well as achieve greater equity in outcomes.

The collection of evidence presented in the previous section suggests that there are several 'low-hanging' fruit for education policy that can improve learning outcomes at low cost. Since the majority of disadvantaged children (especially in rural India) still attend government-run schools, the focus of this section is on the policy priorities that are most relevant to the

[19] Another source of a demand-side market failure can be the high discount rate of parents who may choose to not send their children to school because the benefits are too far in the future while the costs (both monetary and opportunity costs) are immediate. While the Right to Education Act seeks to limit this concern by making schooling compulsory till age 14, there may still be a role for demand-side interventions such as conditional cash transfers at later ages. However, we do not discuss this topic here because (a) the focus of this piece is on primary education, and (b) there is not much good evidence on the impact of conditional transfer programs in India.

running of the government-school system. The paper makes four main policy recommendations in this regard (from easiest to most challenging in terms of practical implementation as well as political feasibility). Implementation issues are discussed in the next section.

4.1. Make Learning Outcomes an Explicit Goal of Primary Education Policy

The evidence on the key role of learning outcomes for both components of the "inclusive growth" agenda of the Government of India combined with the evidence on low levels and trajectories of learning presented in section 2.2, should make it almost obvious that a key goal of primary education policy in India should be to measure and improve learning outcomes.

Nevertheless, this seemingly obvious point is necessary to highlight because the current education policy framework pays almost no attention to it. Nowhere is this more visible than in the "Results Framework Document (RFD)" of the Ministry of Human Resource Development (MHRD). The RFD serves as the document that outlines the goals of MHRD for the year, and places weights on different priorities including access, equity, quality, and departmental processes. While these are all important goals to aspire towards, it is striking that there is *no mention of learning outcomes in the most recent RFD* for 2012-13.[20] While 'quality' of education is given prominence, the document defines quality exclusively in terms of improving the 'inputs' into education – with most of the focus being on teacher training.

This formulation is consistent with standard input-based conceptions of quality of education, but has almost no support in the data. In particular, there is *no study* that finds a positive correlation between a teacher possessing a formal teacher training credential and measures of gains in learning of students taught by the teacher. This is not to suggest that teacher training and other inputs *cannot* be contributors to improving learning outcomes – but to highlight that these inputs *in their current form* do not seem to matter for improved learning outcomes. However, since there is no reason to think that the current policy framework envisages anything other than expanding training and other inputs in their current form, the evidence points to expecting that the future will not be very different from the past experiences.

[20] http://mhrd.gov.in/sites/upload_files/mhrd/files/Modified%20RFD%202012-13_after%20ATF%20meeting.pdf. The closest component of the RFD that relates to learning outcomes is "Assessment of Learners Under Saakshar Bharat" - but this is an adult education scheme.

Of course, there is no guarantee that measuring learning outcomes will by itself lead to an improvement (for instance, 6 years of ASER reports showing consistently low levels of learning have not led to any noticeable changes in policy). But it is almost certain that not measuring outcomes will encourage the system to continue on its current course with poor transformation of inputs into outcomes. Several studies have documented that organizations (especially bureaucracies) are more likely to deliver on outcomes that get measured (Wilson 1989). India's own experience in education over the past decade supports this point, since there has been a significant improvement in input-based measures of quality (which were the stated policy goals). Thus, the starting point in the education policy agenda needs to be an inclusion of improving learning outcomes as an explicit goal of primary education policy with immediate effect.

Opponents of this view raise four sets of objections to this approach. The first is that frequent testing and measurement makes education stressful for children and is therefore not child-friendly (Raina 2013). A second objection is that the Indian education system is already obsessed with exams and test performance to the exclusion of higher-order thinking and critical reasoning, and that Indian education needs less testing and not more. A third objection is that education is a complicated process involving several sets of actors (including parents and the community) and that the Government cannot be held responsible for outcomes (while it can be held accountable for inputs that it is obligated to provide). Finally, even if the principle of outcome-based monitoring is accepted, there is skepticism regarding its administrative feasibility – with a particular concern being the issue of maintaining integrity of measurement if officials will be monitored on the basis of these measures. Each of these points is addressed below.

The first point is well taken, and it is worth highlighting the difference between assessment *of* learning (which is the normal view of testing), and assessment *for* learning (which is what I have in mind). The former approach emphasizes the role of 'testing' what a student knows with a view to ranking and classification (and is inevitably stressful), whereas the latter approach emphasizes the role of assessments as diagnostic tools to teachers and administrators to measure student 'understanding' of concepts to be followed up with targeted instruction (and additional resources where necessary) to bridge learning gaps at an early stage. The entire point of this approach is not to 'stress' the child but to meaningfully 'care' for the child's learning by paying attention to it. This aspect of measurement is in fact consistent with the "Continuous and

Comprehensive Evaluation (CCE)" framework envisaged by the RtE. The recommendation therefore is simply to take this more seriously and require the measurement and reporting of individual student-level learning outcomes over time.

The second objection is based (in my view) on extrapolating the experiences of children in elite high-pressure urban settings (which are the settings experienced by the children of those in policy-making roles) to the entire country. Theory and evidence suggest that optimal policy is different at different levels of learning (see Lazear 2006 for a clear illustration of the relevant issues), and while it is true that excessive testing can narrow the intellectual development of high-achieving students, the opposite is true at low levels of learning (especially given the default policy of automatic promotion through grades regardless of levels of learning). In a setting where 60% of school-aged children cannot read, the evidence suggests that basic and higher order skills are complements and not substitutes (see Muralidharan and Sundararaman 2011b). Further, there is also evidence to suggest that testing helps with processing learned materials and even in the learning of untested materials (Chan et al. 2006). Finally, there is also evidence that parents of rural children (especially those who are not literate themselves) would like to have more objective measures of how their children are doing in school (Andrabi et al 2012). The evidence therefore points to there being *too little* reliable measurement of learning in rural government schools as opposed to too much measurement.

The third objection sounds reasonable but goes completely against the spirit of the Right to Education Act, which places the responsibility of ensuring that every child obtains a quality basic education on the State. If education quality depends on actual learning outcomes as opposed to simply spending time in school, then a natural corollary of the Right to Education Act is that the State takes some responsibility for providing learning skills to all children. Of course, outcomes cannot be guaranteed, but at the very least, measuring and documenting learning levels and gaps provides a basis for differential targeting of additional resources to disadvantaged children to bridge these gaps. Finally, while administrative concerns are very real, these exist with the implementation of almost any policy and different administrative structures can be experimented with at the state and district levels to provide feasible templates for implementation (see section 5.2 for more discussion of this point).

4.2. Consider curricular reform to adjust for the vast variation in learning levels and/or provide additional instructional resources in early schooling years to disadvantaged children with a view to bridging learning gaps at an early age

Muralidharan and Zieleniak (2013) show that the learning trajectories of students over time are substantially flatter than the rate of growth envisaged by the curriculum. It is therefore not surprising that a very large fraction of school-aged children complete primary education without having achieved even basic levels of learning. They also show that there is not only a large amount of variation in student learning levels at the end of grade 1, but that this variance grows over time.

The hypothesis that is most consistent with these findings is one articulated in Chapter 4 of Banerjee and Duflo (2011) and also in Pritchett and Beatty (2012), which is that the curriculum has been designed by highly educated elites and reflects a period of time when there was no expectation of universal primary education. Indeed, as they note, the historical purpose of education systems in many developing countries may not have been to provide 'human capital' to all students as much as to screen gifted students for positions of responsibility in the state and the clergy. Since the teachers continue to follow the textbook as the default mode of instruction, and define their goals in terms of completing the curriculum over the course of year, it is not surprising that they are effectively 'teaching to the top' of the distribution and that a large number of children are in the class but not learning because the lesson is too advanced for them.

While there is no direct test of this hypothesis in the Indian context, it is consistent with the findings of a large body of experimental evaluations of education interventions in India in the past decade. In particular, the finding that targeted remedial instruction programs have been highly effective in improving test scores in spite of being implemented by untrained and poorly paid volunteers, while large investments in teacher qualifications and training, PTR reductions, and other investments in school infrastructure have not been found to be effective suggest that the 'business as usual' pedagogy is not conducive to improving learning outcomes effectively.[21]

[21] This view is also consistent with evidence from multiple studies in Africa. Glewwe, Kremer, and Moulin (2009) provide experimental evidence on the impact of a program that provided free textbooks to children in Kenya. They find that the program had no impact on average test scores, but students at the top 20% of the baseline test score distribution did significantly better with the textbooks. This would clearly make sense if it was only the top 20% of

A natural implication of this theory is that there may be large returns to reforming curricula to move at a different pace for students of different levels (Banerjee and Duflo 2011), or perhaps to even slowing down the pace of the general curriculum (Pritchett and Beatty 2012). However, modifying curricula is a time-consuming and arduous process and waiting to do this, could risk the educational experiences of children in the coming years at a time when there is a very narrow time window left for India's "Demographic Dividend". Thus, while curricular reform to account for variation in learning levels should be a high priority, it may make sense to start immediately with programs that provide supplemental remedial instruction to children who are falling behind in early grades (who would be identified early though a system of CCE as mentioned above).

Banerjee et al (2012) experiment with different models of incorporating learning materials targeted to the initial levels of children into the regular schooling system in Bihar and Uttaranchal. They find that the only model that was successful was one where the instruction was provided in a summer camp, and conclude that the behavior of teachers in the classroom appears to be so deeply ingrained towards completing the 'regular' curriculum that it is difficult for them to deviate from that and modify their behavior towards incorporating the new materials in the classroom.

Thus considerable additional work needs to be done to pilot and evaluate effective models of modifying pedagogy to reflect the need to cater to students who are falling behind. There is however already enough evidence to warrant the scaling up with public funds of programs that provide *supplemental* remedial instruction to children who need it through either after-school programs or through summer camps. The exact implementation models should be left to individual states to determine with the lessons from existing models and evaluations made available to them (see section 5.2 for more on this).

---

students who could read well enough to benefit from possessing a textbook. Duflo, Dupas, and Kremer (2011) present evidence from a program in Kenya that compared test score growth of students in the regular classroom to those of students who were tracked according to initial learning levels. They find that students in the tracked classrooms do significantly better at all initial levels of learning suggesting that reducing the variance of learning levels in the classroom allowed teachers to target the level of the instruction much more effectively

4.3. Expand the use of locally-hired contract teachers (especially for remedial instruction)

The perception that contract teachers are of inferior quality and that their use is a stop-gap measure to be eliminated by raising education spending enough to hire regular teachers is deeply embedded in the status quo education policy discourse (and has been formalized in the RtE). The results discussed in this paper suggest that this view is not supported by the evidence. The fact that all the remedial instruction programs evaluated in this paper used young local volunteers (typically women) who were not trained as teachers and had only a 12th standard qualification (or in some cases even 10th), suggests that motivation and using appropriate pedagogy may be more important determinants of teacher effectiveness than qualifications or training. The results on contract teachers suggest the same conclusion (especially since they are found to be no less effective than regular teachers even with the regular pedagogy).

The combination of low cost, superior performance measures than regular teachers on attendance and teaching activity, and positive overall impact of adding contract teachers to schools suggest that expanding the use of contract teachers could be a highly cost effective way of improving primary education outcomes in India. In particular, expensive policy initiatives to get highly qualified teachers to remote areas (where they are often absent) may be much less cost effective than hiring *several* local contract teachers to provide much more attention to students at a similar cost. Also, as Kingdon and Sipahimalani-Rao (2010) show, there is a surplus of educated unemployed youth (even graduates) who apply for contract and para-teacher jobs even though these jobs pay only a fraction of the salary of a regular teacher. Thus, the supply elasticity of contract teachers appears to be quite high and does not seem to be a binding constraint to expanding the use of locally-hired contract teachers.

The expanded use of contract teachers could address several social challenges at the same time. It would provide employment (and the prestige of a 'white collar' job) to educated unemployed youth, who are not skilled enough for formal sector jobs, but have more than adequate skills to impart basic instruction to first generation learners. Given that the majority of these teachers are young women, the income and autonomy provided by these jobs could improve the intra-household bargaining positions of these women as well as outcomes for their children (as is suggested by many studies). Most important of all, such an initiative could lead to

substantial improvements in learning outcomes of school-aged children - especially if several contract teachers are hired for the cost of one regular teacher.

Opponents of the use of contract teachers worry that their expanded use may lead to a permanent second-class citizenry of contract teachers, which in the long-run will erode the professional spirit of teaching and shift the composition of the teacher stock away from trained teachers towards untrained teachers. Thus, even if expanding the use of contract teachers is beneficial in the short run, it might be difficult to sustain a two-tier system of teachers in the long run. Finally, the political economy concern is that hiring larger numbers of contract teachers will lead to demands to be regularized into civil-service status, which may be politically difficult to resist given the strengths of teacher unions and if such regularization were to happen, it would defeat the purpose of hiring a large number of contract teachers in the first place.

One possible course of action is to hire all new teachers as contract teachers at the school-level, and create a system to measure their performance over a period of time (six to eight years for example) that would include inputs from parents, senior teachers, and measures of value addition using independent data on student performance. These measures of performance could be used in the contract-renewal decision at the end of each fixed-term contract (or to pay bonuses), and consistently high-performing contract teachers could be promoted to regular civil-service rank at the end of a fixed period of time (see the next section for more details). In other words, contract teachers need not be like permanent adjunct faculty, but can be part of a performance-linked tenure track. Continuous training and professional development could be a natural component of this career progression, and integrating contract and regular teachers into a career path should help to address most of the concerns above, including the political economy ones. The recommendation for a career ladder is also made by Kingdon and Sipahimalani-Rao (2010), and by Pritchett and Murgai (2007), who also provide an excellent discussion of how such a system may be implemented in practice.[22]

[22] Pritchett and Murgai (2007) discuss how such a structured career leader for teachers can be embedded within a more decentralized education system that provides local communities more autonomy on managing schools.

4.4. Invest in governance – especially teacher performance measurement and management

Research over the past decade in the US confirms what is intuitive to most observers of education, which is that the most important determinant of education quality that is in the locus of control of policy-makers is teacher quality (Rivkin, Hanushek, and Kain 2005; Rockoff 2004). Good teachers can really make a difference, and a sequence of good teachers can significantly alter the educational trajectory of students and often make up for socio-economic disadvantages (Hanushek and Rivkin 2006). Thus the good news is that education policy makers can have a substantial impact on learning outcomes by hiring and retaining good teachers.

The less good news is that teacher quality as measured by value-addition (which is a statistical measure of the extent to which a teacher is able to improve student learning during the period of time that they are responsible for teaching the concerned student) cannot be predicted by most observable characteristics of teachers (including the factors that are commonly considered to be proxies for quality such as experience, education, and training). Thus, the factors that are rewarded in the status quo may not be the ones that matter for teacher quality. While research on teacher value-added using Indian data is still in early stages, Kingdon and Teal (2010) find very similar results, and preliminary results using the longitudinal data from the APREStt project suggest that the same patterns hold in India.

These results suggest that a better way to identify effective teachers may be to directly measure their value-addition on a regular basis. But, before doing this, it is important to ask if these measures of teacher value-addition are just statistical constructs based on test scores, or if they are useful measures of gains in student human capital. A path-breaking recent paper by Chetty, Friedman, and Rockoff (2011) helps answer this question, by doing a long-term follow up of 2.5 million children in the US and linking their adult outcomes to measures of teacher value-added in grades 3 to 8. They find that teacher quality measured by value addition is strongly predictive of adult outcomes including college attendance, quality of college attended, and wages. Teacher quality in school is also positively correlated with social outcomes such as reduced teenage pregnancy and improved quality of neighborhood lived in. A final striking result is that they estimate that a policy that would replace highly ineffective teachers (those in the bottom 5% of the value-addition distribution over a period of time) with an average teacher would increase lifetime income of students by $300,000.

While these long-term results are not replicable in any Indian dataset at present, preliminary analysis using five-years of longitudinal student data in Andhra Pradesh that is matched to teachers, show that the consequence of variation in teacher quality may be even more pronounced in India. In particular, the difference in mean annual value-added between a teacher who is 1 standard deviation below the mean teacher and one who is 1 standard deviation above the mean is considerably larger than the corresponding figure in US data.[23] Thus, teacher performance measurement and management could be especially high-return activities in the Indian context.

There are two ways to improve average teacher quality: the first is to not hire low-quality teachers and to hire and retain high-quality teachers (the selection margin), the second is to design systems that encourage teachers to exert greater effort in a continuous manner – including upgrading their human capital over the course of their career in ways that improve their teaching ability (the effort margin). However, employing the selection margin effectively under the status quo would be very difficult since the existing selection criteria (especially teacher training) do a very poor job of predicting teacher quality. Thus it is necessary to measure teacher effectiveness on the job before being able to effectively assess their quality.

A career-ladder of the sort proposed in the previous section, whereby all new teachers are hired as contract teachers, provided small annual bonuses on the basis of annual measures of performance, and are then promoted to regular teacher status at the end of a period of time that is long enough to evaluate their performance accurately, would have the dual advantage of improving teacher quality on both the selection as well as the effort margin. Such an initiative could also build a foundation for treating teaching as a true profession where highly effective teachers are rewarded, recognized, and promoted into positions of leadership and mentoring; while ineffective teachers are identified early for coaching and support (and if they are unable to improve even with such support, counseled into other jobs that they may be better suited for). Further details of how such a ladder might work are provided in Pritchett and Murgai (2007).

[23] The exact figures are not quoted here since the results are preliminary, but the inter-quartile range of the teacher value-added distribution in the APRESt data is so much larger than those in US data that the main point is likely to be robust to any changes in the point estimates. Note that a simple explanation for this may be that teachers play a disproportionately large role in test-score gains in a context where many parents are illiterate.

Finally, while putting in place such a system will take time and experimentation to refine the implementation details (see next section), the evidence suggests that even modest investments in better governance can have large returns. A case for optimism in the finding that increased frequency of inspection is correlated with a significant reduction in teacher absence (Muralidharan et al 2013) is that these represent 'business as usual' inspections as currently done by the system. Of course, these are not experimental estimates of the effect of increasing inspections, but the very robust findings of negative correlations between increased inspections and lower absence, suggests that even at the margins of the *current* system, increasing the frequency of supervisory visits to schools is likely to be a more cost effective way of increasing effective teacher-student contact time than hiring more teachers (as seen earlier).

5. Moving from Recommendations to Implementation

While the research to date suggests the four policy recommendations made here, it does not provide adequate guidance as to a possible implementation roadmap. There is perhaps no better proof of the primacy of the implementation challenge than the fact that many of the policy recommendations made in this paper (especially that of a career ladder) are similar to those made five years ago in Pritchett and Murgai (2007) in this *same forum*. There is now more and better evidence to support these recommendations, but the issues have not changed much in the past 5 years and have been clearly visible to experts in this area. The ASER reports have been saying essentially the same thing for seven years now – that learning levels are low in spite of high enrollments – but not much has changed in India's national education priorities (as starkly illustrated by an RFD that has no mention of learning outcomes). The rest of this section outlines some of the key themes that may be relevant to being able to implement an education reform agenda along the lines suggested here.

5.1. Ideas Matter

Even before discussing issues of practical implementation, and political economy, it is worth admitting that the status quo as represented by the formulations in the RtE suggest that the insights from the careful empirical work done on education in India over the past decade using large-scale datasets and paying attention to identification issues, have either not been communicated to or not been accepted by the education 'establishment' in India. To the extent

that the reform agenda being suggested by the quantitative research on the economics of education is seeking to reform the "conventional" wisdom on input-based policies, it is worth thinking about where this conventional wisdom gets formed. At present, it comes from Schools of Education (and related disciplines) where there is a limited amount of quantitative training of students, and where there is a greater emphasis on the history and philosophy of education and of the role of education in shaping society.

These are very important issues, but it has meant that the discourse in education schools and in the 'Education for All (EFA)' and 'Right to Education (RTE)' communities has focused on historical injustices in education access and has typically (and probably correctly) interpreted the lack of universal primary education in India as a failing of the state, representing, at best, elite apathy towards mass education, and at worst an elite conspiracy to make sure that their educational advantage was maintained over generations. Attempts by the "Rights Community" to secure more opportunities for the disadvantaged naturally focus on the most visible symbols of inequity including school buildings, and teachers, which in turn leads to an input-based approach being the default demand of those seeking to secure the rights of disadvantaged children.

Attempts by education economists to bring cost-effectiveness into the discourse are then strongly resisted as an attempt by elites to defund public schools at a time when their own children have all moved to private schools. For instance, one reaction in an education ministry meeting where we presented evidence that locally-hired volunteers and contract teachers may be as (or more) effective than regular trained teachers was that "this will be used by the finance ministry to cut the budget for education". So perhaps one way to bring cost-effectiveness into the conversation is to assure education advocates that the total funding will not be cut even if more cost-effective policy options are followed, and *that any resulting savings will be used to improve education outcomes further*. Of course, the setting of annual departmental budgets is a deeply political process, but such a commitment can serve as a starting point in moving the conversation from "how can we maximize the budgetary allocation for education" to "how can we maximize the quality of education delivered at any given budget" – with an assurance that being efficient will not hurt the sector's budget allocation.

More broadly, active attempts need to be made to disseminate and discuss the insights from the quantitative research over the past ten years with members of the education community

and to incorporate some of the tools and methods of modern quantitative research into curricula and syllabi of education schools, so that their graduates are better equipped to engage with this research and its findings. This is a long-term project, but is an important investment in building dialogue and engagement with regards to priorities for education policy across stakeholders from an 'education' perspective and those from a 'cost-effectiveness/public finance' perspective.

5.2. Allow states more autonomy to experiment and innovate with reform ideas

Even those who agree in principle with the recommendations here would (reasonably) worry about the feasibility of implementing such reforms. While they might seem promising theoretically and be supported by the evidence, there is still no guarantee that these reforms might succeed in practice. But implementation is a tactical and administrative issue that needs to account for local conditions and it would therefore be optimal to give states (and even districts) a substantial amount of autonomy with respect to how they may implement the ideas above. In addition to autonomy with regards to implementation of specific initiatives, it would also make sense to give states more autonomy with respect to how they may use their education budgets to best achieve learning goals.

It is therefore a matter for concern that the RtE in its current form mandates uniformity across a broad range of criteria including detailed specifications for building codes and playgrounds, pupil-teacher ratios, teacher qualifications, and teacher salaries. While these norms may be well intentioned and have the goal of raising education in all states to a minimum standard, there are two problems with this approach. The first problem, which is a conceptual one, is that mandating these norms across the country magnifies the risk of making well-intentioned mistakes - because the jurisdiction over which the mistake is being made would be all of India (which is the largest education system in the world). The second problem, which is an empirical one, is that these are *all* input-based standards, and *none* of these inputs appear to matter much for learning outcomes. Even if experts at the Central-level were to feel that input-based standards are a good starting point for improving education quality, both theory and evidence from other contexts suggests that a better approach would be for the Central Government to issue *guidelines* on suggested inputs (as opposed to mandates) and targets on

outcomes, but then allow states to take the lead in innovating with respect to ways of achieving these outcomes.[24]

Using states as laboratories for education policy innovation makes sense for several reasons.[25] The first is simply that this provides 28 settings for experimentation as opposed to just one, allowing a greater diversity of ideas and implementation models to be tried out at lower risk. Second, Indian states are large (the 10 most populous Indian states would each rank in the top 25 countries in the world by population) and have enough scale to be autonomous policy-making entities on almost all issues related to primary education. Third, there is great diversity among states' political leaders, and corresponding variation in their priorities and their abilities to build political support for specific education policies - which is likely to result in a broader range of ideas being tried. Finally, the locus of political accountability is increasingly shifting to the states, which provides an incentive for states to copy good ideas from each other.[26]

A more productive role for the Central Government would be to support experimentation by states to better understand the impacts of specific initiatives in assessment, pedagogy, resource use, and governance and to then facilitate knowledge transfers across states that enable scaling up of successful reforms. Under the suggested framework for centre-state relations, the Centre would not be looking to institute mandates and police the fulfilling of individual line items, but rather to look to learn from state-level experiences in achieving improvements in learning outcomes, and play a facilitating role in evaluating and transferring knowledge about best practices.[27] This would also be consistent with the first principles of the optimal allocation of roles across levels of government in a federal structure, which suggest that functions having more economies of scale should reside in higher levels of government, whereas those that need

[24] Of course, there is a trade-off here as well, and it may be important for the Central-government to reserve the right to intervene in the cases of states that are not making adequate progress in achieving universal education goals. Nevertheless, the importance of experimentation with solutions and customization of solutions to local contexts suggests an overall approach of centrally-determined minimum goals on education outcomes, with considerable autonomy to states on how to achieve these goals.

[25] This paragraph is based on Muralidharan (2011).

[26] A good example of this is the wide imitation of the Government of Bihar's program to provide bicycles to girls entering secondary school.

[27] An example where such an approach would have been useful is the case of Tamilnadu shifting to a system of Activity-Based Learning (ABL) that features mixed age classrooms and organizing students by learning levels. In principle, the idea of ABL addresses some of the key pedagogical challenges of dealing with variation in learning levels that we discussed earlier. But ABL was rolled out across the Tamilnadu with very little evaluation of the impact of this state-wide change in pedagogy on learning outcomes, which was a missed opportunity for other states (and also for Tamilnadu) to learn more about the impact of this change.

to respond to local information and variation in local conditions should reside in lower levels of government (see Pritchett and Pande 2006 for further discussions on this theme).

In his public remarks at the release of the most recent ASER report in January 2012, the Honourable Minister for HRD, Shri Kapil Sibal, said that Pratham should take the message to Chief Ministers and engage with them to improve outcomes. This is exactly the right approach – but needs to be accompanied with more autonomy for states, untied funds for innovation, and more structured sharing of best practices across states. Even states might be too large a unit for making comprehensive changes quickly, and the appropriate administrative unit for experimenting with some of these ideas may be a district. In fact, a promising approach may be for a committed NGO that can bring the requisite expertise together to work in partnership with an interested state government at the level of one district (or perhaps one district each in a few states) to bring about systemic changes across the district by following the recommendations laid out here. This should be accompanied by careful evaluations of both processes and outcomes to allow comparison of the status quo and the suggested reforms to subject these reform ideas to rigorous testing and evaluation.[28]

5.3. Political Economy – Bringing Teachers on Board

Naturally, many of the reforms outlined here – especially those relating to use of contract teachers – can be expected to be met with opposition from teachers and unions. Nevertheless, it is also true that many teachers are not satisfied with the status quo (as documented in Pritchett and Murgai 2007). This view is supported in the data on teacher absence: Kremer et al (2005) show that in Indian government schools, teachers reporting high levels of job satisfaction are *more likely* to be absent. In subsequent focus group discussions with teachers, it was suggested that this was because teachers who were able to get by with low effort were quite satisfied, while hard-working teachers were dissatisfied because there was no difference in professional outcomes between them and those who shirked. In such a context, the provision of even small amounts of bonuses based on objective measures of performance that are transparently and fairly applied could *increase* intrinsic motivation, and teacher satisfaction, which may lead to teachers

[28] This is something that Pratham is already doing as seen in the results presented in Banerjee et al (2012), but is something that can be considered and attempted more – especially by the larger non-profits that have dedicated endowment-based funding, which will allow them to make longer term investments in personnel and capacity needed to support governments in pilots for 'systemic' transformation.

favoring such a system. It could also explain how average bonuses of only 3% of annual pay could elicit the teacher responses that led to large gains in student learning outcomes in the APRESt experiment.

Muralidharan and Sundararaman (2011a) analyze teacher opinions on performance-linked pay and find that over 80% of teachers had a favorable opinion about the idea of linking a component of pay to measures of performance with over 45% of teachers having a *very favorable* opinion. Over 75% of teachers report an increase in motivation as a result of the program and 68% responded that the government should scale up the program implemented in Andhra Pradesh. Finally, when asked about their preferences over a series of mean-preserving spreads of pay based on performance, 75% of teachers reported support for at least a small portion of pay being linked to performance. What is especially interesting is that levels of teacher support for performance-pay in all these questions were significantly higher in the treatment groups than in the control groups – and thus exposure to a well-designed and communicated program increased teacher support for the idea.

Of course, the opinions of individual teachers could differ from those of teachers as a group and those of union leaders who would wield a disproportionate influence in policy conversations.[29] But, these results suggest that a well-structured career ladder based on objective measures of teacher performance supplemented by inputs from parents and community members may be implementable – especially if total compensation for existing teachers goes up as a result.

More broadly, it is essential for conversations on education reform to bring teachers on board and avoid an adversarial framing of the sort implied by discussions of 'teacher accountability'. Rather, it is important to highlight that all high-performing organizations have well defined goals and feature personnel policies that reward and recognize strong performers. Thus, reforms that improve measurement of learning outcomes, promote effective school leadership and management, and create career rewards for high-performing teachers are likely to increase the professionalism of the education system and increase the respect accorded to the teaching profession.

[29] Unions have a strong history of being against attempts to differentiate pay on the basis of productivity (Ehrenberg and Schwarz 1986)

6. Conclusion

This paper has provided a summary of the insights from a decade of high-quality empirical research on primary education in India and seeks to help bridge the gap between what we are learning from this research and the status quo of primary education policy in India.

The combination of ASER data over time and the international benchmarks provided by the latest PISA results unambiguously establish that the Indian primary schooling system is not doing an adequate job in preparing the generation of children that represents India's "Demographic Dividend" with even the basic skills that will enable them to participate in the process of India's economic growth. The research summarized in this paper highlights that simply increasing the inputs to primary education in a 'business as usual' way are unlikely to change the trajectories of student learning in a meaningful way unless accompanied by significant changes in pedagogy and/or improvements in governance.

The reform agenda suggested in this paper includes some ambitious components. One is the suggestion for re-evaluating the entire curriculum to see if the pace at which the school syllabus is expected to move is a feasible one for all children and to see if slowing down the curriculum and/or introducing some kind of tracking might make sense. The other is to take teacher performance measurement and management seriously. Both of these will take time to figure out the details for and the prudent approach would be to consider serious experiments at the district (or even block) level before trying to implement these ideas on a larger scale.

But there are also items in the list of recommendations that can be done more immediately. For instance, given what we now know about the low levels of learning, it is unconscionable to not make improving learning outcomes a central objective of education policy in India – a good start would be to give it prominence in the "Results Framework Document (RFD)" of MHRD. The good news is that given the (relatively) positive track record of the Indian state in making headway on numbers that are actively monitored – this step alone may catalyze creative thinking in states and districts on ways to improve indicators on learning levels. The research also strongly supports scaling up supplemental instruction programs using locally-hired short-term teaching assistants that are targeted to the level of learning of the child – which should be more easily implementable.

The best approach for implementing this reform agenda would be for the Central Government under the 12th Plan to prioritize learning outcomes and provide states with pools of flexible funding that will allow them to experiment with ways of improving learning outcomes in a cost-effective way. The Planning Commission can help in knowledge-sharing by convening state education departments and providing them with summaries of relevant research; guidelines on what the research points to as effective ways of improving learning outcomes; and in working with states and other partners to design, implement, and evaluate district (or block) level pilots in re-orienting pedagogy and governance towards a better functioning education system.

The next ten years will see the largest ever number of citizens in the Indian school system at any point in the country's history (or future), and it is critical that this generation that represents the demographic dividend be equipped with the literacy, numeracy, and skills needed to participate fully in a rapidly modernizing world. In a fiscally-constrained environment, it is also imperative to use evidence to implement *cost-effective* policies that maximize the social returns on any given level of public investment. The growing body of high-quality research on primary education in the past decade provides an opportunity to put this principle into practice.

References:

Accountability Initiative. 2012. PAISA Report. New Delhi: Center for Policy Research.

Afridi, Farzana. 2010a. "Child Welfare Programs and Child Nutrition: Evidence from a Mandated School Meal Program in India." *Journal of Development Economics* no. 92:152-165.

———. 2010b. The Impact of School Meals on School Participation: Evidence from Rural India. Indian Statistical Institute.

Afridi, Farzana, Bidisha Barooah, and Rohini Somanathan. 2010. School Meals and Student Participation in Urban India. Delhi School of Economics.

Andrabi, Tahir, Jishnu Das, and Asim Khwaja. 2012. Report Cards: The Impact of Providing School and Child Test-scores on Educational Markets. Harvard Kennedy School.

Atherton, Paul, and Geeta Kingdon. 2010. The relative effectiveness and costs of contract and regular teachers in India. Institute of Education, University of London.

Banerjee, Abhijit, Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani. 2010. "Pitfalls of Participatory Programs: Evidence From a Randomized Evaluation in Education in India." *American Economic Journal: Economic Policy* no. 2 (1):1-30.

Banerjee, Abhijit, Rukmini Banerji, Esther Duflo, and Michael Walton. 2012. Effective Pedagogies and a Resistant Education System: Experimental Evidence on Interventions to Improve Basic Skills in Rural India. MIT.

Banerjee, Abhijit, Shawn Cole, Esther Duflo, and Leigh Linden. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *Quarterly Journal of Economics* no. 122 (3):1235-1264.

Banerjee, Abhijit, and Esther Duflo. 2011. *Poor Economics*: MIT Press.

Barro, Robert J. 1991. "Economic Growth in a Cross Section of Countries." *Quarterly Journal of Economics* no. 106 (2):407-43.

Benhabib, Jess, and Mark M Spiegel. 1994. "The Role of Human Capital in Development: Evidence from Aggregate Country Data." *Journal of Monetary Economics* no. 34 (2):143-74.

Borkum, Evan, Fang He, and Leigh Linden. 2010. School Libraries and Language Skills in Indian Primary Schools: A Randomized Evaluation of the Akshara Library Program. Columbia.

Chan, Jason C. K. , Kathleen B. McDermott, and Henry L. Roediger III. 2006. "Retrieval-Induced Facilitation: Initially Nontested Material Can Benefit From Prior Testing of Related Material." *Journal of Experimental Psychology: General* no. 135 (4):553-571.

Chetty, Raj, John N Friedman, and Jonah E. Rockoff. 2011. The Long-Term Impact of Teachers: Teacher Value-Added and Student Outcomes in Adulthood. Harvard.

Chin, Aimee. 2005. "Can Redistributing Teachers Across Schools Raise Educational Attainment? Evidence from Operation Blackboard in India." *Journal of Development Economics* no. 78:384-405.

Cristia, Julian P, Pablo Ibarraran, Santiago Cueto, Ana Santiago, and Eugenio Severin. 2012. Technology and Child Development: Evidence from the One Laptop Per Child Program. IZA Discussion Paper 6401.

Das, Jishnu, Stefan Dercon, James Habyarimana, Pramila Krishnan, Karthik Muralidharan, and Venkatesh Sundararaman. 2013. "School Inputs, Household Substitution, and Test Scores." *American Economic Journal: Applied Economics* no. 5 (2):29-57.

Das, Jishnu, and Tristan Zajonc. 2010. "India shining and Bharat drowning: Comparing two Indian states to the worldwide distribution in mathematics achievement." *Journal of Development Economics* no. 92:175-187.
Desai, Sonalde, Amaresh Dubey, Reeve Vanneman, and Rukmini Banerji. 2009. "Private Schooling in India: A New Educational Landscape." In *India Policy Forum*, edited by Suman Bery, Barry Bosworth and Arvind Panagariya. New Delhi: Sage.
Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *The American Economic Review* no. 91 (4):795-813.
Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review* no. 101 (5):1739-74.
———. 2012. School Governance, Teacher Incentives, and Pupil-Teacher Ratios: Experimental Evidence from Kenyan Primary Schools. National Bureau of Economic Research. Working Paper 17939.
Duflo, Esther, Rema Hanna, and Stephen Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review* no. 102 (4):1241-78.
Duraisamy, P. 2002. "Changes in Return to Education in India, 1983 - 94: By Gender, Age-cohort and Location." *Economics of Education Review* no. 21 (609-622).
Educational Initiatives. 2010. School Learning Study. Educational Initiatives.
Ehrenberg, Ronald G, and Joshua L Schwarz. 1986. "Public-Sector Labor Markets." In *Handbook of Labor Economics*, edited by Orley Ashenfelter and Richard Layard. Elsiever.
Glewwe, Paul, Michael Kremer, and Sylvie Moulin. 2009. "Many Children Left Behind? Textbooks and Test Scores in Kenya." *American Economic Journal: Applied Economics* no. 1 (1):112-135.
Goyal, Sangeeta, and Priyanka Pandey. 2011. "Contract teachers in India." *Education Economics*:1-21.
Hanushek, Eric, and Steve Rivkin. 2006. "Teacher Quality." In *Handbook of the Economics of Education*, edited by Eric Hanushek and Finis Welch. North-Holland.
Hanushek, Eric, and Ludger Woessmann. 2008. "The Role of Cognitive Skills in Economic Development." *Journal of Economic Literature* no. 46 (3):607-668.
———. 2010. Do Better Schools Lead to More Growth: Cognitive Skills, Economic Outcomes, and Causation. CESIfo, Munich.
He, Fang, Leigh Linden, and Margaret Macleod. 2008. Teaching What Teachers Don't Know: An Assessment of the Pratham English Language Program. UT Austin.
Jacob, Verghese, Anjini Kochar, and Suresh Reddy. 2008. School Size and Schooling Inequalities. Stanford.
Jayaraman, Rajshri, Dora Simroth, and Francis De Vericourt. 2010. The Impact of School Lunches on Primary School Enrollment: Evidence from India's Mid-Day Meal Scheme. Indian Statistical Institute.
Jensen, Robert. 2010. "The (Perceived) Returns to Education and the Demand for Schooling." *Quarterly Journal of Economics*.
———. 2012. "Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India." *Quarterly Journal of Economics* no. 127 (2):753-792.

Kingdon, Geeta Gandhi, and Vandana Sipahimalani-Rao. 2010. "Para-Teachers in India: Status and Impact." *Economic and Political Weekly* no. XLV (12):59-67.
Kingdon, Geeta, and Francis Teal. 2010. "Teacher Unions, Teacher Pay and Student Performance in India: A Pupil Fixed Effects Approach." *Journal of Development Economics* no. 91 (2):278-288.
Kremer, Michael, Karthik Muralidharan, Nazmul Chaudhury, F. Halsey Rogers, and Jeffrey Hammer. 2005. "Teacher Absence in India: A Snapshot." *Journal of the European Economic Association* no. 3 (2-3):658-67.
Kumar, Krishna, Manisha Priyam, and Sadhna Saxena. 2005. "The trouble with para-teachers." *Frontline* no. 18 (22).
Lakshminarayana, Rashmi, Alex Eble, Preetha Bhakta, Chris Frost, Peter Boone, Diana Elbourne, and Vera Mann. 2012. Support to Rural India's Public Education System: the STRIPES Cluster Randomised Trial of Supplementary Teaching, Learning Material and Additional Material Support in Primary Schools.
Lazear, Edward. 2006. "Speeding, Terrorism, and Teaching to the Test." *Quarterly Journal of Economics* no. 121 (3):1029-1061.
Linden, Leigh. 2008. Complement or Substitute? The Effect of Technology on Student Achievement in India. Columbia.
Malamud, Ofer, and Christian Pop-Eleches. 2011. "Home Computer Use and the Development of Human Capital." *Quarterly Journal of Economics* no. 126 (2):987-1027.
Mankiw, Gregory, David Romer, and David Weil. 1992. "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics* no. 107 (May):407-437.
Mukerji, Shobhini, and Michael Walton. 2012. Learning the Right Lessons: Measurement, Experimentation and the need to turn India's Right to Education Act upside down. Harvard Kennedy School.
Munshi, Kaivan;, and Mark R. Rosenzweig. 2006. "Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy." *American Economic Review* no. 96 (4):1225-1252.
Muralidharan, Karthik. 2011. "India's States Can be Laboratories for Policy Innovation." *Business Standard*, December 12
———. 2012. Long Term Effects of Teacher Performance Pay: Experimental Evidence from India. UC San Diego.
Muralidharan, Karthik, Jishnu Das, Alaka Holla, Michael Kremer, and Aakash Mohpal. 2013. The Fiscal Costs of Weak Governance: Evidence from Teacher Absence in India. UC San Diego.
Muralidharan, Karthik, and Michael Kremer. 2008. "Public and Private Schools in Rural India." In *School Choice International*, edited by Paul Peterson and Rajashri Chakrabarti. Cambridge: MIT.
Muralidharan, Karthik, and Nishith Prakash. 2013. Cycling to School: Increasing Secondary School Enrollment for Girls in India. UC San Diego.
Muralidharan, Karthik, and Venkatesh Sundararaman. 2010. "The Impact of Diagnostic Feedback to Teachers on Student Learning: Experimental Evidence from India." *Economic Journal* no. 120 (546):F187-F203.
———. 2011a. "Teacher Opinions on Performance Pay: Evidence from India." *Economics of Education Review* no. 30 (3):394-403.

———. 2011b. "Teacher Performance Pay: Experimental Evidence from India." *Journal of Political Economy* no. 119 (1):39-77.
———. 2013. Contract Teachers: Experimental Evidence from India. UC San Diego.
Muralidharan, Karthik, and Yendrick Zieleniak. 2013. Meauring Learning Trajectories in Developing Countries with Longitudinal Data and Item Response Theory. UC San Diego.
Nguyen, Trang. 2008. Information, Role Models, and Perceived Returns to Education: Experimental Evidence from Madagascar. MIT.
Olken, Ben. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* no. 115 (2):200-249.
Pratham. 2012. *Annual Status of Education Report.*
Pritchett, Lant. 2012. The first PISA results for India: The end of the beginning. edited by Ajay Shah.
Pritchett, Lant, and Amanda Beatty. 2012. The negative consequences of over-ambitious curricular in developing countries. Harvard Kennedy School.
Pritchett, Lant, and Rinku Murgai. 2007. "Teacher Compensation: Can Decentralization to Local Bodies Take India From Perfect Storm Through Troubled Waters to Clear Sailing?" In *India Policy Forum*, edited by Suman Bery, Barry Bosworth and Arvind Panagariya. Sage Publications.
Pritchett, Lant, and Varad Pande. 2006. Making Primary Education Work for India's Rural Poor: A Proposal for Effective Decentralization. In *Social Development Papers: South Asia Series*. New Delhi: World Bank.
PROBE Team. 1999. *Public Report on Basic Education in India*. New Delhi: Oxford University Press.
Raina, Vinod. 2013. Tests are Torture. *India Today*, January 25.
Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* no. 73 (2):417-58.
Rockoff, Jonah E. 2004. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *American Economic Review* no. 94 (2):247-252.
Schoellman, Todd. 2012. "Education Quality and Development Accounting." *Review of Economic Studies* no. 79:388-417.
Shastry, Kartini. 2012. "Human Capital Response to Globalization: Education and Information Technology in India." *Journal of Human Resources* no. 47 (2):287-330.
Shields, Robin. 2011. "ICT or I see tea? Modernity, technology and education in Nepal." *Globalization, Societies, and Education* no. 9 (1):85-97.
Wilson, James Q. 1989. *Bureaucracy*. New York: Basic Books.