Better Teachers, Better Results?
Evidence from rural Pakistan

Marine de Talancé

This is an Accepted Manuscript of an article published by Taylor & Francis in The Journal of Development Studies on 20 Dec 2016, available at:
https://www.tandfonline.com/doi/abs/10.1080/00220388.2016.1265944?journalCode=fjds20

Most of the existing literature examining the determinants of school quality in developing countries has failed to take into account the crucial role of teachers. This study assesses how teachers contribute to knowledge acquisition in Punjab, Pakistan. The baseline specification used is a gain model with three different levels of fixed effects. We find that teacher quality is strongly correlated with student achievement. Increasing teachers' wages could improve schooling quality, as could the recruitment of local and contract teachers. Our analysis also underlines the importance of reforming training programs and re-thinking wage policies.

Keywords: Education, Pakistan, Primary Schooling, Quality, Skills, Teachers, Temporary contracts.

JEL classification: I21, I25, I28.

1. Introduction

There is widespread evidence on the poor quality of primary schooling in developing countries, where many pupils leave primary school without basic mathematical and reading skills (Muralidharan & Sundararaman, 2013). This poor quality has substantial economic consequences. When education is of low quality, students may be incited to drop out of school sooner (Hanushek, Lavy, & Hitomi, 2008) leading to less human-capital accumulation. This reduces individual future earnings and exacerbates inequalities (Behrman, Ross, & Sabot, 2008; Boissiere, Knight, & Sabot, 1985). From a macroeconomic point of view, this translates into lower economic growth (Hanushek & Woessmann, 2012).

Whereas research in developed countries has shown that teachers have a considerable impact on student achievements (Behrman et al., 2008; Chetty, Friedman, & Rockoff, 2014; Hanushek, 2011), most work in developing countries has not included teachers in the appropriate education function. The limited progress towards understanding the impact of teachers on academic achievement in developing countries mainly reflects data limitations.

In this paper, we ask to what extent student achievement is linked to teachers. We make use of a unique panel dataset on third-, fourth- and fifth-grade students in three districts of Punjab province in Pakistan. Over the course of three years, these children, in both private and public schools, were tested in three different subjects (Mathematics, Urdu and English). This project also gathered rich information on households, schools and teachers.

We use the panel dimension of the data to develop an appropriate identification strategy. This has three main key features distinguishing it from prior work in developing countries. First, we estimate a gain model, where we take into account the effect of prior knowledge on current achievement. Second, we exploit variation in scores within schools and teachers to control for diverse aspects of selection. Third, in order to control for unobserved student heterogeneity, we also use student fixed effects, comparing students who were assigned to teachers with different characteristics over time. Our results show that teacher quality is

strongly correlated with student achievement. Increasing teachers' wages could improve schooling quality, as could the recruitment of local and contract teachers. Recruiting local and female teachers could reduce gender inequalities of academic achievement. Our analysis also suggests that policy reforms concerning training programs and the design of wages should be implemented. These findings are robust to different specifications, score measurements and sample restrictions.

The remainder of this paper is structured as follows. Section 2 reviews the relevant literature on teachers and student achievement, and Section 3 the educational context in Pakistan. Section 4 discusses the empirical methodology. In Section 5, we describe the database and the variables used, and the results appear in Section 6. Finally, the last section concludes with implications for educational policies in Pakistan and further research.

2. Related literature

Since the influential Coleman report (Coleman et al., 1966), international evidence has shown that traditional input-based policies have failed to improve the quality of education (see Glewwe & Kremer, 2006; Glewwe, Hanushek, Humpage, & Ravina, 2011; Murnane & Ganimian, 2014 for reviews of the literature). In the absence of natural or randomized experiments, various papers have appealed to education-production functions. However, until recently, most of these functions did not take into account the effects of teachers on child performance.

Recent work in developed countries has included teacher fixed effects in education-production functions. Hanushek and Rivkin (2010) review ten recent studies in the United States and show that, on average, a one standard-deviation rise in teacher effectiveness raises students' reading and mathematics scores by respectively 0.13 and 0.17 of a standard deviation. These results are confirmed by papers linking teacher effectiveness to students' future earnings (Behrman et al., 2008; Chetty et al., 2014; Hanushek, 2011).

However, when specific teacher characteristics are included in education production functions, the results are not convincing as very few observables explain the differences in learning (Hanushek, 2003; Glewwe et al., 2011).

Aslam and Kingdon (2011) use data on 65 schools in Lahore district, Pakistan. They find no evidence that observable teacher characteristics affect student achievement. Fehrler, Michaelowa, and Wechtler (2009) estimate education-production functions in 21 Sub-Saharan countries from the SACMEQ and PASEC databases. They conclude that teacher education and professional training do not affect student achievement as they do not reflect teacher knowledge. Michaelowa (2001) also uses the PASEC database for five African countries and finds that teacher job satisfaction is positively associated with student learning. Aturupane, Glewwe, and Wisniewski (2013) estimate fourth-grade students' academic performance in Sri Lanka including a small number of teacher characteristics (teacher experience and number of meetings with parents). These teacher variables are insignificant in instrumental-variable estimations. García Palomer and Paredes (2006) use Chilean data and find that observable teacher practices explain only a small part of student learning. Araujo, Carneiro, Cruz-Aguayo, and Schady (2016) measure the impact of teachers on kindergarten students in Ecuador and find that teacher behaviors are strongly associated with gains in learning. Andrabi, Bau, Das, and Khwaja (2010) use the same database as that analyzed here to evaluate the impact of private schools on learning. While their results are interesting, they do not control for student past performance and do not specifically focus on teachers. Das and Bau (2014) make use of the same database to look at the relationship between teacher pay and productivity, using the rise in contract teachers as a natural experiment. Their empirical method is different as they estimate teacher value-added (teacher fixed effects) without children fixed effects. They then regress teacher value-added on teacher characteristics and compare the public and private sectors. They do not directly relate student outcomes to teacher characteristics.

While it is therefore generally acknowledged that teacher quality is key in improving

education, to our knowledge no paper has convincingly determined what lies behind teacher effectiveness. Previous work in developing countries has been hampered by a lack of panel data, and has not been able to take into account the dynamic dimension of learning and address student selection.

3. Quality of schooling and teachers in Pakistan

3.1 Education in Pakistan

While 10% of primary age out-of-school children in the world live in Pakistan (UNESCO, 2014), many indicators suggest that there have been educational improvements over the last decades. Between 1971 and 2012, gross enrollment ratios increased from 47% to 93%, from 16% to 36% and from 2% to 9% for respectively the primary, secondary and tertiary levels. Nevertheless, with a literacy rate of 54.9%, Pakistan compares poorly to the average in South Asia. Of the children attending primary schools, 39% drop out before completing the last grade. The educational system in Pakistan continues to discriminate against girls and children from poor households and rural areas (DHS, 2013).

Even when they do go to school, Pakistani children do not necessarily learn the basics. According to the ASER national survey, by the end of primary school only 55% of children can read a story in Urdu, Sindhi or Pashto, 49% a sentence in English, and 50% can divide 2-digit numbers (ASER, 2015).

3.2 Teachers in Pakistan

Andrabi, Das, Khwaja, Vishwanath, and Zajonc (2011) dedicate a whole chapter to teacher quality in Pakistan, using the same database as we do here. The poor quality of learning in schools in Pakistan is often attributed to teachers (Westbrook et al., 2009). This could notably reflect the limited qualification requirements (10 years of education) for becoming a

primary-school teacher. The professional training of teachers in Pakistan is not standardized or based on acceptable standards. Although the National Education Policy (2009) states that a Bachelor degree in Education (B.Ed.)[1] should be the minimum required to teach at the elementary level, Primary Teaching Certificate (PTC) or the Certificate in Teaching (CT) certifications[2] remain dominant (NEMIS-AEPAM, 2011). Few teachers benefit from continuous training programs, as these remain voluntary.

The recruitment of teachers is also problematic, being based on political pressure and not merit (Ali, 2000). Prior to 1997, teachers in Punjab were mostly hired as permanent public servants, but this led to politically-motivated recruitment and transfers preventing the most competent teachers from entering the system. In 1997, a ban on hiring new teachers was implemented to deal with a budgetary crisis. In 2002, the ban was removed and teachers were increasingly hired on five-year renewable contracts. This growth of contract teachers is similar to the situation in India (Muralidharan & Sundararaman, 2013) and other countries in Sub-Saharan Africa (Bourdon, Frölich, & Michaelowa, 2010). The efficiency of this recruitment policy remains uncertain. Contract teachers may exert more effort in order to have their contracts renewed. However, contract teachers are less qualified and less well-trained.

The teaching profession has for decades had little appeal and low social status in Pakistan, and is often perceived as the last choice for young professionals (Westbrook et al., 2009; UNESCO, 2013). However, teachers in Pakistan are on average as affluent as other individuals with at least 10 years of education and teachers' salaries have recently risen (UNESCO, 2013).

4. Model

The model developed in this section is based on Chetty, et al. (2014), Meghir and Rivkin (2011) and Todd and Wolpin (2003). The achievement of child i at the end of the school year t, A_{it}, is a function of the entire history of family inputs (H_{i0}, H_{i1}, ..., H_{it}), child

characteristics (C_{i0}, C_{i1}, ..., C_{it}), school inputs (QS_{i0}, QS_{i1}, ..., QS_{it}) and the child's initial endowment (μ_{i0}).

$$A_{it} = a_t(H_{i0}, H_{i1}, ..., H_{it}, C_{i0}, C_{i1}, ..., C_{it}, QS_{i0}, QS_{i1}, ..., QS_{it}, \mu_{i0}) \quad (1)$$

Following Todd and Wolpin (2003), endowed ability is assumed to be invariant. Assuming that past human capital depreciates at a constant rate ($1-\delta$), the cognitive skill of child i in grade t, A_{it}, can be represented as the depreciated knowledge at grade $t-1$ plus the investment made in t, I_{it}.[3]

$$A_{it} = \delta A_{i,t-1} + I_{it} \Leftrightarrow A_{it} = \sum_{j=0}^{t} \delta^j I_{t-j} \quad (2)$$

Here δ is the persistence coefficient.

The investment in grade t is represented by a reduced-form as a function of all the inputs in the period under consideration from the family (H_{it}), the child (C_{it}) and the school (QS_{it}), plus the effect of innate ability (μ_{i0}):

$$I_{it} = \beta_t H_{it} + \alpha_t C_{it} + \gamma_t QS_{it} + \zeta_t \mu_{i0} + e_{it} \quad (3)$$

The impact of these inputs over time decays according to both the distance in time between the investment and the output, and the grade when the investment was made (which is why there is a subscript *t* on the coefficients). As equation (3) is linear, its substitution into equation (2) yields:

$$A_{it} = \sum_{j=0}^{t} \delta^{t-j} [\beta_j H_{ij} + \alpha_j C_{ij} + \gamma_j QS_{ij} + \zeta_j \mu_{i0} + \varepsilon_{ij}] \quad (4)$$

Where QS_{i0} (the school investment in year zero) is null as the child has not yet attended

school.

The econometric estimation of equation (4) is problematic, as the genetic endowment (μ_{i0}) is unobserved and there are no datasets with all of the past and current inputs. Further assumptions are required in order to lighten the historical data requirements.

The simplest specification, known as the contemporaneous model, assumes that there is an immediate and complete decay of previous knowledge ($\delta = 0$) or that inputs do not vary over time (current inputs capture all of the history of inputs). This is the approach used by Dee (2004) and Rockoff (2004).

$$A_{it} = \beta H_{it} + \alpha C_{it} + \gamma QS_{it} + \varepsilon'_{it} \tag{5}$$

The error term includes the child's endowment ($\varepsilon'_{it} = \zeta_t \mu_{i0} + \varepsilon_{it}$). We require another assumption to estimate equation (5) correctly: contemporaneous inputs must be uncorrelated with unobserved innate ability. These two assumptions are unrealistic. The skills acquired in previous periods are likely to persist over time (Todd & Wolpin, 2003; Cunha, Heckman, & Schennach, 2010; Meghir & Rivkin, 2011). If parents react to child endowments by investing more in more "gifted" children, then the second assumption does not hold and equation (5) is inconsistent with OLS estimation. A fixed-effects model can be used here, but this does not solve the question of omitted past inputs and the model remains biased. For these reasons, value-added models are generally preferred (Todd & Wolpin, 2003).

The restricted value-added model, also known as the gain score model, assumes that there is perfect persistence of past knowledge ($\delta = 1$). Various contributions in the literature have used gain models (Hanushek, Kain, O'Brien, & Rivkin, 2005; Harris & Sass, 2011; Rivkin, Hanushek, & Kain, 2005).

$$A_{it} - A_{it-1} = \beta H_{it} + \alpha C_{it} + \gamma QS_{it} + \varepsilon''_{it} \quad (6)$$

For this empirical specification to be valid, inputs have to be uncorrelated with the error term in t and the impact of each input (and of innate ability) must be independent of the grade when they were applied. The error terms must also be serially correlated and match the rate of decay of the inputs.

The third model, known as the unrestricted value-added model, makes no assumptions about the value of persistence δ.

$$A_{it} = \delta A_{it-1} + \beta H_{it} + \alpha C_{it} + \gamma QS_{it} + \varepsilon'''_{it} \quad (7)$$

Equation (7) will be consistent if the effects of all inputs (including child innate ability) are assumed to decline at the same geometric rate. Estimating the value-added model is still problematic, as prior achievement is by construction correlated with the child's endowment which is captured by the error term. Because of data limitations, many studies have ignored the correlation between lagged achievement and the error (Aaronson, Barrow, & Sander, 2007; Clotfelter, Ladd, & Vigdor, 2007). As using fixed effects in a dynamic model introduces bias (Nickell, 1981), the lagged variable needs to be instrumented using the generalized method of moments (Andrabi et al., 2011).

Although this model is less restrictive than the gain model, due to the smaller temporal dimension of the database, as in Andrabi et al. (2011) the Hansen tests reject the hypothesis of instrument validity. The gain model is our preferred model (equation (6)), while the unrestricted value-added (equation (7)) and contemporaneous (equation (5)) models are used for robustness tests.

The vector of school inputs, QS_{it}, contains school-level inputs, denoted by S_{imt} with m indexing the school, a vector of classroom inputs (peer characteristics and educational

material), P_{it}, and teacher characteristics, T_{ijt}, with j indexing the teacher. The value-added model without fixed effects is:

$$A_{it} = \delta A_{it-1} + \beta H_{it} + \alpha C_{it} + \gamma T_{ijt} + \phi S_{imt} + \eta P_{it} + \varepsilon_{it}''' \quad (8)$$

Assuming that school-level inputs are constant over the time span of the survey, they can be captured using school fixed effects.[4] This approach reduces the bias associated with students and teachers sorting into schools (Harris & Sass, 2011).

$$A_{it} = \delta A_{it-1} + \beta H_{it} + \alpha C_{it} + \gamma T_{ijt} + \eta P_{it} + s_m + \varepsilon_{it}''' \quad (9)$$

Where s_m represents the school fixed effects.

To control for individual heterogeneity, and especially innate ability, children fixed effects can also be included. Although individual fixed effects are not recommended in an unrestricted value-added model, student fixed effects can be added to the gain model, which yields:

$$A_{it} - A_{it-1} = \beta {H_{it}}^{v} + \alpha {C_{it}}^{v} + \gamma T_{ijt} + \eta P_{it} + s_m + c_i + e_{it} \quad (10)$$

Where H_{it}^{v} and C_{it}^{v} refer to time-varying child and family characteristics. The invariant family and individual inputs are captured by the individual fixed effects, c_i.

Teacher characteristics can also be estimated through fixed effects, τ_j, yielding:

$$A_{it} - A_{it-1} = \beta {H_{it}}^{v} + \alpha {C_{it}}^{v} + \gamma\, {T_{ijt}}^{v} + \eta P_{it} + \tau_j + s_m + c_i + e'_{it} \quad (10)$$

Where ${T_{ijt}}^{v}$ represents time-varying teacher characteristics. Note that here s_m are

school-year fixed effects in order to avoid multicollinearity with teacher fixed effects.

5. Data and variables

5.1 Description of the data

The data come from the Learning and Educational Achievement in Pakistan Schools (LEAPS) project. Between 2001 and 2005, the LEAPS project collected data on the distribution of schools in rural Punjab and the quality of education. Children were tested in three subjects: Mathematics, Urdu (the vernacular) and English.

The sample covers 823 schools (in the first round) in 112 villages in three districts: Attock (North), Faisalabad (Central) and Rahim Yar Khan (South). The sample is not nationally representative as villages were randomly chosen from a list of villages with both public and private schools. As expected, these villages are wealthier, larger and more educated than the average rural village.[5] All private and public schools within the village boundary and within a short walk of any village household were surveyed.[6] Multiple questionnaires were distributed to different groups (school principals, teachers and children) to obtain a complete picture of the educational environment.

During the first round, children in grade three were tested. They were tracked and retested in grades four and five. Among those who were tested during the first wave, 87% were retested in the second or third waves and 67% were tested at all waves.

For the purpose of this study, the original sample of students was reduced to 15,470 children after keeping schools with at least two different teachers and after dropping children who repeated the grade or advanced two grades at once. As the use of teacher fixed effects requires students to have different teachers, the estimates of equation (11) rely on the 70% of children who change teachers at least once over the span of the survey. A probit model is estimated to analyze whether these children have particular characteristics. We observe

almost no significant background differences once teacher and school characteristics are controlled for. [7]

5.2 Econometric issues

Student attrition is relatively low, as 87% of first wave children were tested at least twice. A probit model to assess whether children who attrit have particular characteristics shows that only few observable variables predict attrition (Online Appendix, Table A1.1). In the fixed-effects model, attrition is only problematic if the selection is correlated with idiosyncratic errors (Wooldridge, 2010). Following Verbeek and Nijman (1992), if the idiosyncratic errors are uncorrelated with attrition, the lead selection variable indicating attrition in $t+1$ should not affect achievement in t. The results suggest that attrition is not related to idiosyncratic errors (Online Appendix, Table A1.2). Given that student attrition is relatively low, and that it is not correlated with idiosyncratic errors, the estimates using school and student fixed effects are unlikely to be biased by student attrition. The results using the balanced student sample are nonetheless provided as a robustness check.

The estimated effects of observed teacher characteristics will also be biased if teacher attrition is not accounted for (Rockoff, 2004; Kane, Rockoff, & Staiger, 2008; Harris & Sass, 2011). Around 15% of teachers left their schools over the previous two years, and most of them did so due to personal and family issues, with only 17% joining another school. A probit model of teacher attrition shows that this could bias our estimates (Online Appendix, Table A1.3). Following Hanushek et al. (2005); Harris and Sass (2011) and Rockoff (2004), we thus provide estimates including both student and teacher fixed effects. However, if unobserved time-varying teacher characteristics are correlated with the probability of attrition, they will not be adequately captured by teacher fixed effects (Harris & Sass, 2011). We will hence also provide estimates from the balanced teacher sample.

If students, school resources and teachers are not assigned to schools and classrooms

randomly, the estimates could be biased (Ishii & Rivkin, 2009). Despite the relatively large number of schools, school choice is mainly driven by distance and budgetary constraints and not by the quality of the school (Online Appendix, Tables A2.1 and A2.2), especially as parents appear to have little information on school quality (Online Appendix, Table A2.3). Non-random assignment is unlikely to come from the schools themselves, as they admit almost all children who apply.[8] The inclusion of school fixed effects partially deals with the non-random time-invariant assignment of students to schools. Within a school, if students are assigned to specific teachers, the estimates could also be biased. Following a two-step procedure implemented by Rockoff (2004), we test for systematic classroom assignment and find no evidence of the systematic matching of students to teachers (Online Appendix, Table C2.4). Moreover, the inclusion of child fixed effects alleviates this bias. The only source of potential bias comes from dynamic student matching to teachers (Rothstein, 2010; Koedel & Betts, 2011).

We could also have bias from teachers non-randomly selecting into contracts. In a model with teacher fixed effects, biases exist if teacher contracts are correlated with time-varying unobservable characteristics such as productivity. A probit analysis of teacher contract suggests that teacher selection into contracts does not depend on performance.[9] This selection into contracts is thus not a major source of bias as it depends on variables for which we control.

Finally, we test whether the effects of the different inputs are constant over time, and whether child's past achievements affect current inputs. The results support the gain model (Online Appendix, Tables A3.1 and A3.2).

5.3 Variables

The scores are calculated using the Item Response Theory (IRT) method, which is widely used in educational assessments such as PISA or TIMMS. Contrary to Classic Test Theory, IRT gives

different weights to correct answers depending on the difficulty of the question.[10] The scores are standardized by year and subject.

The descriptive statistics appear in Table 1. Child health is measured using the World Health Organization Reference 2007 as body mass by age for children aged between five and nineteen. Children with z-scores more than two standard deviations below the reference population are considered malnourished, whereas those with z-scores more than two standard deviations above the reference are considered overweight.

Table 1: Descriptive Statistics

	Panel sample		First wave		Second wave		Third wave	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Child Characteristics								
Child age	10.29	1.53	9.64	1.48	10.49	1.49	11.16	1.53
Girl	0.46	0.50	0.45	0.50	0.44	0.50	0.44	0.50
Child overweight	0.02	0.14	0.01	0.12	0.02	0.13	0.03	0.17
Child underweight	0.14	0.34	0.12	0.32	0.13	0.34	0.14	0.35
Household Characteristics								
Dad not educated	0.33	0.47	0.35	0.48	0.34	0.48	0.33	0.47
Dad less primary	0.06	0.23	0.07	0.26	0.07	0.25	0.06	0.24
Dad primary to high sec	0.54	0.50	0.51	0.50	0.52	0.50	0.53	0.50
Dad over high sec	0.07	0.25	0.06	0.25	0.07	0.26	0.08	0.26
Mum not educated	0.64	0.48	0.65	0.48	0.64	0.48	0.63	0.48
Mum less primary	0.07	0.25	0.07	0.25	0.07	0.26	0.07	0.26
Mum primary to high sec	0.28	0.45	0.26	0.44	0.27	0.44	0.28	0.45
Mum over high sec	0.01	0.11	0.01	0.12	0.01	0.12	0.01	0.12
Wealth index	0.03	1.49	-0.33	1.39	-0.10	1.45	0.32	1.53
Teacher Characteristics								
Female teacher	0.56	0.50	0.56	0.50	0.54	0.50	0.52	0.50
Same gender teacher	0.83	0.37	0.82	0.39	0.84	0.36	0.84	0.37
Local teacher	0.62	0.49	0.39	0.49	0.72	0.45	0.73	0.44

Teacher's years of exp	11.14	8.62	9.92	7.93	11.37	8.92	12.19	9.09
PTC training	0.44	0.50	0.47	0.50	0.45	0.50	0.40	0.49
CT training	0.19	0.39	0.18	0.38	0.18	0.38	0.21	0.41
BED training	0.14	0.34	0.12	0.32	0.13	0.34	0.18	0.38
Teacher's years of education	11.51	1.64	11.49	2.11	11.45	1.38	11.64	1.25
Non-permanent contract teacher	0.38	0.49	0.38	0.48	0.38	0.49	0.40	0.49
Teaching wage (Rupees)	4845	2847	4166	2274	4808	2675	5697	3194
Can receive a bonus	0.32	0.47	0.35	0.48	0.35	0.48	0.26	0.44
Did receive bonus or prize	0.09	0.28	0.09	0.29	0.09	0.28	0.07	0.26
Teacher other work: agriculture	0.15	0.36	0.17	0.38	0.12	0.32	0.15	0.36
Teacher other work: business	0.02	0.15	0.03	0.17	0.02	0.15	0.02	0.15
Teacher other work: teaches outside	0.13	0.33	0.11	0.31	0.13	0.34	0.14	0.35
Teacher other work	0.03	0.17	0.03	0.16	0.03	0.18	0.03	0.16
Teacher absence (no. of days last month)	1.95	2.70	2.15	3.20	1.83	2.60	1.93	2.21
Classroom Characteristics								
Class size	28.62	18.20	28.54	16.44	31.66	20.53	27.70	17.52
% with English books	0.87	0.21	0.77	0.29	0.90	0.15	0.92	0.13
% with Maths books	0.87	0.21	0.78	0.28	0.90	0.16	0.92	0.13
% with Urdu books	0.87	0.21	0.79	0.28	0.90	0.16	0.91	0.14
% with desks	0.54	0.45	0.46	0.44	0.55	0.46	0.60	0.46
% with chairs	0.19	0.38	0.16	0.36	0.19	0.38	0.21	0.39
% with blackboards	0.84	0.29	0.82	0.31	0.84	0.28	0.85	0.27
% girls in the class	0.46	0.42	0.45	0.41	0.44	0.43	0.44	0.42
District Characteristics								
Attock	0.33	0.47	0.34	0.47	0.35	0.48	0.33	0.47
Faisalabad	0.37	0.48	0.35	0.48	0.35	0.48	0.34	0.47
Rahim Yar Khan	0.30	0.46	0.31	0.46	0.30	0.46	0.33	0.47
Observations	20565		11553		11765		10367	
No. of children	6855		11553		11765		10367	

Notes: PTC and CT refer respectively to the Primary Teaching Certificate and the Certificate in Teaching. BED refers to a Bachelor degree in Education.

Source: Author, using the three waves of the LEAPS data.

We construct a wealth index following Filmer and Pritchett (2001) by applying a Principal Component Analysis to household asset indicators.[11] Local teachers work in the village in which they were born. Since the 2002 reform, teachers have increasingly been hired under temporary contracts.[12] To test whether training programs and experience are collinear, we estimate a probit for teacher training as a function of teacher characteristics. Teachers with a Bachelors degree in Education are not significantly less experienced.[13] On average, teachers earn 4,854 Rs ($ 46.2) per month, although there is great heterogeneity between teachers.

6. Empirical results

6.1 Student achievement

The three-way fixed effects estimates allow us to evaluate teacher effects on student achievement (Table 2). The F-statistics for the joint significance of the teacher fixed effects (p-values below 0.001) show that teachers strongly predict achievement in all three subjects. Table 2 shows the raw standard deviations of the teacher fixed effects: these standard deviations overestimate teacher effects due to sampling error (Rockoff, 2004; Aaronson et al., 2007). Following Aaronson et al. (2007), we adjust the variance of the teacher fixed effects by subtracting the average sampling variance estimated as the mean of the square of the standard errors of the estimated teacher fixed effects. A one standard deviation rise in the distribution of teacher fixed effects increases scores by 0.6 to 1 of a standard deviation. These estimates are higher than those in developed countries. Hanushek and Rivkin (2010) review 10 studies in the United States, where the estimates vary between 0.08 and 0.36. Our larger effect size may be explained by the greater variation in teacher quality in low-income countries like Pakistan (Das & Bau, 2014). Such comparisons are however problematic as one standard deviation in the United States may well not be comparable to that in Pakistan.

Table 2: Regression with Teacher Fixed Effects

Dependent variable: scores in

	English	Maths	Urdu
Variations in teacher fixed effects			
Raw standard deviation	0.689	1.026	0.679
Adjusted standard deviation	0.635	0.984	0.626
No. of teacher fixed effects estimates	339	344	352
No. of observations	11268	11268	11268
Adjusted R^2	0.545	0.582	0.539
Student FE	Yes	Yes	Yes
School-by-year FE	Yes	Yes	Yes
Lagged scores	Yes	Yes	Yes
Student time-varying covariates	Yes	Yes	Yes

Table 3 presents the estimates of the gain model (equation (6)). This model relates the differences in achievement gains to differences in teacher and child characteristics. There are no significant differences in achievement between boys and girls. Not surprisingly, older children in wealthier households perform significantly better in all subjects. Parental education has little or no impact on student achievement, but is probably partly captured by wealth.

Table 3: Estimates of the gain model

Dependent variable: IRT gain score - ML	(1) English	(2) Maths	(3) Urdu	(4) English	(5) Maths	(6) Urdu
Girl	0.037*	0.011	0.031			
	(0.021)	(0.022)	(0.023)			
Child age	0.020***	0.015***	0.012**	0.015	-0.008	0.004
	(0.006)	(0.006)	(0.006)	(0.056)	(0.058)	(0.057)
Child underweight	0.001	0.032	0.023	0.027	-0.067	0.063
	(0.022)	(0.023)	(0.023)	(0.120)	(0.125)	(0.131)
Child overweight	0.029	-0.070	-0.032			
	(0.048)	(0.050)	(0.051)			
Dad less primary	0.009	0.030	-0.008			
	(0.028)	(0.030)	(0.032)			

Dad primary to high sec	0.012	0.016	-0.007			
	(0.015)	(0.015)	(0.015)			
Dad over high sec	0.022	0.080***	-0.013			
	(0.028)	(0.027)	(0.027)			
Mum less primary	-0.016	0.025	-0.024			
	(0.025)	(0.025)	(0.027)			
Mum primary to high sec	-0.027*	0.010	-0.004			
	(0.016)	(0.016)	(0.016)			
Mum over high sec	0.026	-0.060	-0.017			
	(0.042)	(0.047)	(0.046)			
Wealth index	0.009*	0.019***	0.018***	0.060***	0.089***	0.098***
	(0.005)	(0.006)	(0.006)	(0.020)	(0.021)	(0.021)
Female teacher	-0.092	-0.206**	-0.193**	-0.107	-0.257***	-0.192*
	(0.080)	(0.081)	(0.086)	(0.098)	(0.097)	(0.098)
Same gender teacher	-0.005	-0.009	-0.028	-0.187**	-0.063	-0.088
	(0.022)	(0.023)	(0.024)	(0.090)	(0.088)	(0.088)
Local teacher	0.121***	0.142***	0.107**	0.110**	0.166***	0.064
	(0.045)	(0.047)	(0.046)	(0.053)	(0.056)	(0.053)
Teacher exp	0.012	0.002	0.018*	0.006	0.000	0.029**
	(0.011)	(0.010)	(0.011)	(0.012)	(0.011)	(0.013)
Teacher exp^2	-0.000	0.000	-0.000	-0.000	-0.000	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Teacher education	0.259**	0.197	0.204	0.206	0.175	0.172
	(0.130)	(0.143)	(0.131)	(0.160)	(0.177)	(0.151)
Teacher education^2	-0.009*	-0.007	-0.008	-0.008	-0.007	-0.007
	(0.005)	(0.006)	(0.005)	(0.007)	(0.008)	(0.006)
Non-permanent contract	0.250***	0.212**	0.252***	0.269***	0.206**	0.269***
	(0.072)	(0.085)	(0.077)	(0.084)	(0.097)	(0.087)
PTC training	0.067	-0.148**	-0.076	0.051	-0.155*	-0.124*
	(0.057)	(0.067)	(0.062)	(0.068)	(0.079)	(0.073)
CT training	-0.124*	-0.104	-0.141*	-0.105	-0.115	-0.206**
	(0.064)	(0.072)	(0.073)	(0.074)	(0.084)	(0.085)
No training	-0.043	-0.129	-0.049	-0.139	-0.150	-0.162*
	(0.079)	(0.095)	(0.081)	(0.095)	(0.116)	(0.097)
Log teacher monthly wage	0.258***	0.230***	0.235***	0.262***	0.227***	0.151**

	(0.064)	(0.059)	(0.060)	(0.079)	(0.073)	(0.073)
Bonus for pupil performance	0.062	-0.038	0.070	0.094	0.015	0.118**
	(0.050)	(0.051)	(0.048)	(0.061)	(0.063)	(0.059)
Bonus for other reasons	0.242**	-0.053	0.085	0.258**	0.056	0.207*
	(0.108)	(0.118)	(0.106)	(0.125)	(0.137)	(0.120)
Teacher absence	-0.000	-0.005	-0.011*	-0.000	-0.003	-0.013**
	(0.007)	(0.006)	(0.006)	(0.007)	(0.007)	(0.006)
Teacher teaches outside	0.005	0.051	0.022	0.049	0.078	0.080
	(0.051)	(0.055)	(0.051)	(0.057)	(0.066)	(0.058)
Teacher other work	0.049	0.014	-0.090**	0.044	0.005	-0.118**
	(0.046)	(0.046)	(0.045)	(0.054)	(0.056)	(0.054)
Class size	-0.010***	-0.017***	-0.012***	-0.008***	-0.015***	-0.010***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
% girls in the class	0.076	0.067	0.114	0.437**	0.315	0.454
	(0.166)	(0.210)	(0.233)	(0.221)	(0.287)	(0.282)
% with English books	-0.304***			-0.250**		
	(0.099)			(0.116)		
% with Maths books		-0.310***			-0.271**	
		(0.111)			(0.133)	
% with Urdu books			-0.421***			-0.433***
			(0.102)			(0.120)
% with desks	0.199***	0.083**	0.102***	0.157***	0.040	0.054
	(0.038)	(0.039)	(0.037)	(0.044)	(0.045)	(0.042)
% with chairs	0.017	0.023	0.070	0.046	-0.008	0.001
	(0.047)	(0.058)	(0.046)	(0.054)	(0.070)	(0.055)
% with blackboards	0.005	-0.031	-0.074	0.020	-0.072	-0.063
	(0.046)	(0.049)	(0.050)	(0.052)	(0.056)	(0.056)
Observations	11181	11181	11181	11181	11181	11181
Adjusted R^2	0.102	0.097	0.078	0.035	0.049	0.048
FE Schools	Yes	Yes	Yes	Yes	Yes	Yes
FE Teachers	No	No	No	No	No	No
FE Students	No	No	No	Yes	Yes	Yes

Notes: Robust clustered standard errors in parentheses: * p<.1, ** p<.05 , *** p<.01

The dependent variables are scores calculated via the Item Response Theory (IRT) method using Maximum Likelihood (ML). PTC and CT refer respectively to the Primary Teaching Certificate and the Certificate in Teaching.

Reference categories: Teacher has followed a BED training program, cannot receive a bonus, does not have another job and has a permanent contract.

Source: Author, using the three waves of the LEAPS data.

Providing more textbooks is less useful than reducing class sizes or providing desks. Smaller class sizes are associated with greater learning gains. The negative estimated coefficients on textbooks may at first appear surprising. However, the relevant literature in developing countries finds no evidence of a significant effect of textbooks on student achievement (Glewwe, Kremer, & Moulin, 2009; Glewwe et al., 2011; Sabarwal, Evans, & Marshak, 2013). There are three main explanations. First, when more textbooks are provided, teachers may be less involved because they think that their lack of involvement will be compensated by the textbooks. Our data confirm this, as teachers spend significantly less time taking breaks when few textbooks are provided.[14] Second, textbooks may be better adapted to stronger students (Glewwe et al., 2009). To test this, the gain model is estimated separately for students who had a low score in *t-1* and the other students. The negative effect of textbooks only holds for students with the lowest initial achievement level.[15] Third, providing more textbooks, but to a small number of students, could increase inequality. When textbooks are introduced in a quadratic form, only English textbooks have a significant impact on students' achievement and this effect is first negative and then positive.[16] The gender composition of the class, as measured by the percentage of girls in the class, positively influences English achievement, probably because girls outperform boys in English.

With respect to teachers, female teachers are associated with significantly worse Mathematics and Urdu scores compared to male teachers. This may be due to a lack of motivation or commitment, as most women teachers declare that they will stop teaching after getting married.[17] Women may also be less involved in their teaching job as they have more household responsibilities. Note that we do however control for teacher absence, which will pick up part of this effect.[18] While there is no obvious benefit from having a same-gender teacher, teacher place of birth does play a large and significant role in explaining learning. Local teachers are more effective, especially in Mathematics. They may speak the same language as the students, which facilitates comprehension. They may also be

of the same caste and share the same values.

Traditional observable teacher characteristics (education and experience) are not associated with better learning. To test for collinearity between experience and contract, we estimate the model without the contract variable: experience remains insignificant.

Compared to teachers with a Bachelors degree in Education, teachers with PTC and CT certifications are less effective. Despite the offer of pre-service and in-service training programs, Pakistani teachers in practice mainly use traditional and conservative teaching methods (Westbrook et al., 2009). Comparing the teaching methods and teacher knowledge of teachers with different types of training, PTC or CT program teachers have less knowledge in English, Mathematics and Urdu compared to teachers with a B.Ed.[19]. Reforming training programs could therefore improve the quality of primary schooling in Pakistan.

Interestingly, students taught by contract teachers outperform those taught by regular teachers. This result is consistent with the literature in South Asia (Atherton & Kingdon, 2010; Muralidharan & Sundararaman, 2013) and Africa (Banerjee, Cole, Duflo, & Linden, 2007; Duflo, Dupas, & Kremer, 2015). However, recruiting more contract teachers could have a negative overall impact if these teachers are less trained, educated and experienced with an effect that outweighs the positive effect of the extra effort they put in (relative to tenured teachers) to make sure that their contracts are renewed. Dropping the training, education, and experience variables continues to produce a positive effect of temporary contracts. Section 7 analyses a number of different plausible explanations of the positive impact of teacher contract.

Teacher pay is positively associated with all student test scores, suggesting that monetary incentives work, as in the efficiency-wage literature and consistent with existing findings in this area (Glewwe et al., 2011; Hanushek, 2003). To test for potential reverse causality (where current pay is determined by past pupil learning outcomes, in a pay for performance sense), we regress the logarithm of current teacher wages on past student scores and other control variables (including school fixed effects). The results, available on request, show that

student past performance is not linked to current teacher wages. Note that this absence of performance-related pay is not in line with recent experimental findings in India (Muralidharan & Sundararaman, 2011) where the linking of teacher performance to wages produced gains in achievement.

Students who have a teacher eligible for bonuses because of good student performance perform better, but only in Urdu and when student fixed effects are included. As shown by Murnane and Ganimian (2014), rewarding teachers for student performance is effective only if these rewards are based on test-scores rather than graduation rates, and we have no indication that this is the case. The effect of bonuses has been shown to be short-lived, as teachers increase their effort to raise short-run test scores by conducting more preparation sessions but do not attend the class more nor change their pedagogical methods (Glewwe, Ilias, & Kremer, 2010). Being eligible for bonuses for other reasons significantly positively affects students' achievement in English.

Teachers who give private tutoring are not significantly more effective. Teachers with another paid job are less effective, at least in Urdu. Because they have other potential sources of income, these teachers do not rely on their teaching job as much as other teachers. Moreover, they are mainly employed in agriculture, an occupation requiring different skills from those needed to teach efficiently. They therefore do not gain useful experience when it comes to teaching. Finally, as expected, teacher absence has a negative impact, but only significantly so in Urdu. The small size of this effect is in line with the existing literature (Aslam & Kingdon, 2010; Michaelowa, 2001). We estimate an OLS regression to understand the drivers of teacher absenteeism.[20] Only four teacher characteristics explain teacher absenteeism: experience, type of contract, training level and gender. The gain model is then re-estimated without these variables, producing very similar results.

Overall then, the main teacher characteristics that positively affect child achievement are salary, geographical background and contract.

6.2 Robustness checks

We performed extensive robustness checks to test our results.[21] First, we have also estimated the gain model including teacher fixed effects in addition to student and school-year fixed effects (equation (11)). Second, instead of IRT subject-specific scores calculated via maximum-likelihood procedures, we use two different score measures: the Classic Test and IRT Expected A Posteriori (EAP) scores. Third, following Harris and Sass (2011), we examine the robustness of our results to changes in the assumed value of the persistence rate δ. The positive effects of temporary-contract teachers, teacher wages and local teachers remain significant when lower persistence rates are assumed, even though the magnitude of their impacts varies. Fourth, as some findings depend on the value of the persistence rates, following Andrabi et al. (2011), we estimate the value-added model (equation (7)) using the difference GMM estimator developed by Arellano and Bond (1991). In line with Andrabi et al. (2011), we find relatively low persistence rates, ranging from 0.09 to 0.46: children lose over half of their achievement in a single year. Fifth, to deal with attrition, we estimate the gain model using the balanced student and teacher panel samples. The previous results are robust to this sample change. All these tests found the results to be virtually identical.

Finally, we also run robustness checks to see if the variables affecting student performance vary by child gender and the type of school attended. The previous results continue to hold (Tables B1 and C2 in the Online Appendix). The benefit from being taught by local and female teachers is more pronounced for girls, suggesting that hiring local and female teachers could reduce the gender gap in academic achievement.

7. The relationships between teacher contract and achievement

One plausible explanation of the strong effect of teacher contracts is that the pressure linked to temporary contracts increases teacher effort and reduces absenteeism (Atherton & Kingdon, 2010; Duflo et al., 2015; Muralidharan & Sundararaman, 2013). This cannot

however be the only explanation, as teacher absenteeism is included in our estimates. Another explanation, put forward by Atherton and Kingdon (2010), is that, because temporary contracts offer lower wages, only individuals who are intrinsically motivated will choose teaching. Moreover, recruiting contract teachers with lower wages allows schools to hire more teachers and reduce the pupil-teacher ratio, which increases student achievement (Atherton & Kingdon, 2010). However, this again cannot be the only explanation as pupil-teacher ratios are controlled for in our empirical analyses.

Teacher effects can be driven by the unobserved characteristics of contracts or differential effects of observable characteristics. To test for the latter, following Atherton and Kingdon (2010), a saturated model is estimated with all of the observed characteristics and their interactions with the contract variable. [22] An F-test of the insignificance of the interaction terms is rejected for all three subjects. Contract teachers mitigate the positive effect of both wages and bonuses on performance confirming that contract teachers are not only motivated by wages (Atherton & Kingdon, 2010). The other significant differential effects reduce the negative effect of class size and increase the positive effect of providing chairs and blackboards.

As temporary contracts are alleged to put more pressure on employees, the impact of contract may depend on the end of this contract. A teacher with a terminating contract faces more pressure than a teacher with a number of contract years remaining. To test this, the gain model is estimated with a dummy indicating for the contract expiring soon (Table 5). Section D in the Online Appendix describes how this variable was created. The effect of contract teachers is partly explained by the pressure the teacher faces to perform well so that their contract is renewed.

Table 4: Gain model – effect of the end of the contract

Dependent variable: IRT gain score – ML	(1) English	(2) Maths	(3) Urdu

Non-permanent contract	0.272***	0.172*	0.236***
	(0.086)	(0.099)	(0.088)
Non permanent contract*expires soon	-0.010	0.128**	0.125**
	(0.054)	(0.065)	(0.059)
Observations	11181	11181	11181
Adjusted R^2	0.035	0.050	0.049
FE Schools	Yes	Yes	Yes
FE Teachers	No	No	No
FE Students	Yes	Yes	Yes

Notes: Robust clustered standard errors in parentheses: * p<.1, ** p<.05 , *** p<.01

The dependent variables are scores calculated using the Item Response Theory (IRT) method via Maximum Likelihood (ML).

Reference categories: Teacher has followed a BED training program, cannot receive a bonus, does not have another job and has a permanent contract. *Control variables*: Same variables as in Table 2.

Source: Author, using the three waves of the LEAPS data.

To compare the cost of a contract teacher to the associated gains in terms of student achievement, we follow Atherton and Kingdon (2010) and estimate the cost per predicted achievement gain point using the gain model with school fixed effects. On average, contract teachers earn less than one-third the salary of their civil-service colleagues (Table 5). However, they are associated with greater academic gains. The predicted gains for students with contract teachers are 2.6 to 4.5 times higher. Therefore, on average, the cost of increasing gains in scores by one point is 9 to 16 times higher for regular teachers.

Table 5: The relative cost of contract teachers

		Regular Teachers	Contract Teachers	Ratio (regular/contract)
Average monthly salary (Rupees)		7000.68	2020.39	3.47
Predicted mean gain in score	English	0.62	2.81	0.22
	Maths	1.06	2.72	0.39

	Urdu	1.10	3.63	0.30
Cost per predicted point	English	11211.15	718.14	15.61
gain in achievement	Maths	6635.30	743.84	8.92
(Rupees)	Urdu	6393.24	557.02	11.48

Notes: The predicted mean score gain is calculated using a gain model with school fixed effects. The predicted mean scores are calculated from a contemporaneous regression with school fixed effects. These predicted values hold all other independent variables at their mean values.

8. Discussion and conclusion

We here use a gain model to analyze the teacher characteristics that affect student skills in primary schools in three districts of the Punjab province in Pakistan. We include both school and student fixed effects to control for the non-random sorting of students and teachers and unobserved heterogeneity. We also provide estimates using teacher fixed effects.

We find strong evidence for a relationship between teachers and skill acquisition. The results suggest that teachers are one of the main drivers of learning. Certain observable teacher characteristics are associated with student achievement: contract teachers perform better than permanent teachers; locally-recruited teachers are more effective; and our results suggest that higher wages may motivate teachers and improve the quality of schooling. All of these policies are easier to implement than traditional policies such as greater teacher education or experience.

The relationship between teachers' wages and students' achievement raises the question of the design of wages. The current wage policy is linked to characteristics that are not associated with teacher efficiency. It is therefore crucial to reexamine the wage policy using the literature on teacher performance pay (Muralidharan & Sundararaman, 2011).

Teacher experience and education have relatively little impact on students' achievement. It would nevertheless be of interest to look separately at total teaching experience and experience in the current school. When teachers stay for a long time in the school, they may

learn teaching methods that are adapted to the particular children in the school (Hanushek, Kain, & Rivkin, 1999; Boyd, Lankford, Loeb, & Wyckoff, 2005). Were the results to confirm this intuition, teacher retention should also be a priority for school principals.

Recruiting local teachers improves the quality of learning, especially in Mathematics. We would like to understand what lies behind this relationship. It may be because teachers are from the same caste as their students, share the same values or speak the regional language. Unfortunately, data limitations here prevent us from exploring further.

Another pathway to schooling quality lies in contract teachers and the pressure they face to have their contracts renewed. Future research should look into the most efficient ways of assessing teacher quality in order to decide whether to renew their contracts. It would also be helpful to know if permanent teachers who are periodically evaluated by supervisors perform as well as contract teachers. If this is the case, two different policies could be implemented: recruiting more contract teachers or increasing the supervision of both permanent and contract teachers.

Last, the reform of teacher-training courses is required to improve the quality of primary education in Pakistan. Further analysis is therefore crucial in order to pinpoint the deficiencies of the current programs.

9. Notes

[1] The Bachelor in Education (B.Ed) is a one-year post-graduate qualification program after a Bachelors degree.

[2] Admission to CT and PTC training programs requires 10 and 12 years of schooling for respectively primary and middle school teachers. After one year of training, teachers receive respectively the Primary Teaching Certificate (PTC) or the Certificate in Teaching (CT).

[3] $A_{it} = \delta A_{i,t-1} + I_{it} = \delta(\delta A_{i,t-2}) + \delta I_{i,t-1} + I_{it} = \delta^3 A_{i,t-3} + \delta^2 I_{i,t-2} + \delta I_{i,t-1} + I_{it} = \sum_{j=0}^{t} \delta^j I_{t-j}$

[4] Data rarely exist on time-varying school inputs. One exception could be school principal characteristics. However, given the time span of the survey we use (three years) and little variation over time in school characteristics, we use school fixed effects.

[5] Note that at the time of the LEAPS survey, around 50 percent of the rural population in Punjab lived in villages with at least one private school (Andrabi, Das, Khwaja, Vishwanath , & Zajonc, 2007). Since 2001, Pakistan has been split up into 148 districts (36 in the Punjab province).

[6] In each selected village, all schools within the village and schools within 15 minutes walking distance for Attock and Faisalabad and 30 minutes for Rahim Yar Khan, a less densely-populated district, were surveyed. Villages with more than 24 schools were excluded.

[7] These results are available on request.

[8] While 81% of children attend a school where there is a particular procedure for admitting pupils, most schools (98%) admit every student

who applies.

[9] These results are available on request.

[10] Two students who answer the same number of items will not be scored identically unless they answered the same set of items correctly.

[11] The asset indicators used to construct a single wealth index are a radio, a TV, a fridge, a motorcycle or a scooter, a car, taxi, van or pickup and a telephone. We exclude agricultural assets on purpose.

[12] Note that before the 2002 reform, some teachers were already hired under temporary contracts. Among the sample teachers who were hired before the 2002-reform, 23% have a temporary contract. The 2002-reform considerably increased this proportion. 95% of sample teachers hired after the 2002 reform have temporary contracts.

[13] These results are available on request..

[14] These results are available on request.

[15] These results are available on request.

[16] These results are available on request.

[17] 27% of the sample female teachers (vs. 16% of male teachers) declare they would stop teaching after getting married and 47% (vs. 3%) declare they would stop depending on their spouse's decision.

[18] Female teachers spend on average five hours doing housework per day (vs. three hours for male teachers).

[19] These results are available on request.

[20] These results are available on request.

[21]These results are available on request.

[22] These results are available on request

10. References

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and Student Achievement in the Chicago Public High Schools. *Journal of Labor Economics, 25*, 95-135. doi:10.1086/508733

Ali, M. A. (2000). Supervision for teacher development: an alternative model for Pakistan. *International Journal of Educational Development, 20*, 177-188. doi: 10.1016/S0738-0593(99)00020-6

Andrabi T., Bau N., Das J., Khwaja A. (2010). *Are Bad Public Schools Public 'Bads?' Test Scores and Civic Values in Public and Private Schools*. (Working Paper No. 11904). Cambridge, MA: Harvard University. Retrieved from: http://epod.cid.harvard.edu/files/epod/files/khwaja_a_-_bad_public_schools_2010-01-20.pdf?m=1449508051

Andrabi, T., Das, J., Khwaja, A. I., & Zajonc, T. (2011). Do Value-Added Estimates Add Value? Accounting for Learning Dynamics. *American Economic Journal: Applied Economics*, *3*(3), 29-54. doi:10.1257/app.3.3.29

Andrabi, T., Das, J., Khwaja, A. I., Vishwanath, T., & Zajonc, T. (2007). *Learning and Educational Achievements in Punjab Schools (LEAPS): Insights to inform the education policy debate*. Washington, DC: World Bank. Retrieved from: http://siteresources.worldbank.org/PAKISTANEXTN/Resources/Publications-and-Reports/367132-1208398596098/2008LEAPS.pdf

Arellano, M., & Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and An Application to Employment Equations. *Review of Economic Studies, 58*, 277-297. doi:10.2307/2297968

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher Quality and Learning Outcomes in Kindergarten. *The Quaterly Journal of Economics, 131*, 1415-1453. doi:10.1093/qje/qjw016

ASER. (2015). *Annual Status of Education Report: ASER-Pakistan 2015*. Lahore: South Asian Forum for Education Development. Retrieved from: http://www.aserpakistan.org/document/aser/2015/reports/national/ASER_National_Report_2015.pdf

Aslam, M., & Kingdon, G. (2011). What can teachers do to raise pupil achievement? *Economics of Education Review, 30*, 559-574. doi:10.1016/j.econedurev.2011.01.001

Atherton, P., & Kingdon, G. (2010). *The relative effectiveness and costs of contract and regular teachers in India*. (Working Paper No. 2010-15). Oxford University: CSAE. Retrieved from: http://2010.economicsofeducation.com/user/pdfsesiones/036.pdf

Aturupane, H., Glewwe, P., & Wisniewski, S. (2013). The impact of school quality, socioeconomic factors, and child health on students' academic performance: evidence from Sri Lankan primary schools. *Education Economics, 21*, 2-37. doi: 10.1080/09645292.2010.511852

Banerjee, A., Cole, S., Duflo, E., & Linden, L. (2007). Remedying Education: Evidence from Two Randomized Experiments in India. *The Quaterly Journal of Economics*, *122*, 1235-1264. doi:10.1162/qjec.122.3.1235

Behrman, J. R., Ross, D., & Sabot, R. (2008). Improving quality versus increasing the quantity of schooling: Estimates of rates of return from rural Pakistan. *Journal of Development Economics*, *85*, 94-104. doi:10.1016/j.jdeveco.2006.07.004

Boissiere, M., Knight, J. B., & Sabot, R. H. (1985). Earnings, Schooling, Ability, and Cognitive Skills. *American Economic Review, 75*, 1016-1030. Retrieved from: https://www.jstor.org/stable/1818642?seq=1#page_scan_tab_contents

Bourdon, J., Frölich, M., & Michaelowa, K. (2010). Teacher shortages, teacher contracts and their effect on education in Africa. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 173, 93-116. Retrieved from: https://www.jstor.org/stable/20622580?seq=1#page_scan_tab_contents

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). Explaining the short careers of high-achieving teachers in schools with low-performing students. *American Economic Review*, *95*(2), 166-171. doi: 10.1257/000282805774669628

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review*, *104*, 2593-2632. doi: 10.1257/aer.104.9.2593

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review, 26*, 673-682. doi:10.1016/j.econedurev.2007.10.002

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. (1966). *Equality of educational opportunity*. Washington, D.C.: Office of Education. Retrieved from: http://files.eric.ed.gov/fulltext/ED012275.pdf

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica, 78*, 883-931. doi: 10.3982/ECTA6551

Das, J., & Bau, N. (2014). *The Misallocation of Pay and Productivity in the Public Sector: Evidence from the Labor Market for Teachers*. Unpublished manuscript.

Dee, T. S. (2004). Teachers, Race, and Student Achievement in a Randomized Experiment. *Review of Economics and Statistics, 86*, 195-210. doi: 10.1162/003465304323023750

DHS. (2013). *Pakistan Demographic and Health Survey 2012-2013.* Islamabad, Pakistan, and Calverton, Maryland, USA: NIPS and ICF International
Retrieved from: http://microdata.worldbank.org/index.php/catalog/1918/get_microdata

Duflo, E., Dupas, P., & Kremer, M. (2015). School governance, teacher incentives, and pupil–teacher ratios: Experimental evidence from Kenyan primary schools. *Journal of Public Economics*, *123*, 92-110. doi: 10.1016/j.jpubeco.2014.11.008

Fehrler, S., Michaelowa, K., & Wechtler, A. (2009). The Effectiveness of Inputs in Primary Education: Insights from Recent Student Surveys for Sub-Saharan Africa. *Journal of Development Studies, 45*, 1545-1578. doi:10.1080/00220380802663625

Filmer, D., & Pritchett, L. (2001). Estimating Wealth Effects without Expenditure Data-or Tears: An Application to Educational Enrollments in States of India. *Demography, 38*, 115-132. Retrieved from: http://www.jstor.org/stable/3088292

Garcia Palomer, C., & Paredes, R. D. (2010). Reducing the Educational Gap: Good Results in Vulnerable Groups. *Journal of Development Studies*, *46*, 535-555. doi:10.1080/00220380903318038

Glewwe, P. W., Hanushek, E. A., Humpage, S. D., & Ravina, R. (2011). *School Resources and Educational Outcomes in Developing Countries: A review of the literature from 1990 to 2010* (Working Paper No. 17554). Cambridge, MA: National Bureau of Economic Research. doi: 10.3386/w17554

Glewwe, P., & Kremer, M. (2006). Schools, Teachers, and Education Outcomes in Developing Countries. *Handbook of the Economics of Education*, *2*, 945-1017. doi: 10.1016/S1574-0692(06)02016-2

Glewwe, P., Kremer, M., & Moulin, S. (2009). Many Children Left Behind? Textbooks and Test Scores in Kenya. *American Economic Journal: Applied Economics*, *1*(1), 112-135. doi:10.1257/app.1.1.112

Glewwe, P., Ilias, N., & Kremer, M. (2010). Teacher Incentives. *American Economic Journal: Applied Economics,* 2(3), 205-227. doi:10.1257/app.2.3.205

Hanushek, E. A. (2003). The Failure of Input-based Schooling Policies. *Economic Journal, 113*, 64-98.

doi:10.1111/1468-0297.00099
Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review, 30*, 466-479. doi:10.1016/j.econedurev.2010.12.006
Hanushek, E. A., & Rivkin, S. G. (2010). Generalizations about Using Value-Added Measures of Teacher Quality. *American Economic Review, 100*(2), 267-271. doi: 10.1257/aer.100.2.267
Hanushek, E. A., & Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth, 17*, 267-321. doi:10.1007/s10887-012-9081-x
Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1999). *Do Higher Salaries Buy Better Teachers?* (Working Paper No. 7082). Cambridge, MA: National Bureau of Economic Research. doi: 10.3386/w7082
Hanushek, E. A., Kain, J. F., O'Brien, D. M., & Rivkin, S. G. (2005). *The Market for Teacher Quality* (Working Paper No. 11154). Cambridge, MA: National Bureau of Economic Research. doi:10.3386/w11154
Hanushek, E. A., Lavy, V., & Hitomi, K. (2008). Do Students Care about School Quality? Determinants of Dropout Behavior in Developing Countries. *Journal of Human Capital*, 2, 69-105. doi: 10.1086/529446
Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics, 95*, 798-812. doi:10.1016/j.jpubeco.2010.11.009
Ishii, J., & Rivkin, S. G. (2009). Impediments to the Estimation of Teacher Value Added. *Education, 4*, 520-536. doi:10.1162/edfp.2009.4.4.520
Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review, 27*, 615-631. doi:10.1016/j.econedurev.2007.05.005
Koedel, C., & Betts, J. R. (2011). Does Student Sorting Invalidate Value-Added Models of Teacher Effectiveness? An Extended Analysis of the Rothstein Critique. *Education, 6*, 18-42. doi:10.1162/EDFP_a_00027
Meghir, C., & Rivkin, S. (2011). Econometric Methods for Research in Education. *Handbook of the Economics of Education*, *3*, 1-87. doi: 10.1016/S0169-7218(11)03001-2
Michaelowa, K. (2001). Primary Education Quality in Francophone Sub-Saharan Africa: Determinants of Learning Achievement and Efficiency Considerations. *World Development, 29*, 1699-1716. doi: 10.1016/S0305-750X(01)00061-4
Muralidharan, K., & Sundararaman, V. (2011). Teacher Performance Pay: Experimental Evidence from India. *Journal of Political Economy, 119*, 39-77. doi:10.1086/659655
Muralidharan, K., & Sundararaman, V. (2013). *Contract Teachers: Experimental Evidence from India* (Working Paper No. 19440). Cambridge, MA: National Bureau of Economic Research. doi: 10.3386/w19440
Murnane, R. J., & Ganimian, A. J. (2014). *Improving Educational Outcomes in Developing Countries: Lessons from Rigorous Impact Evaluations* (Working Paper No. 20284). Cambridge, MA: National Bureau of Economic Research. doi: 10.3386/w20284
NEMIS-AEPAM. (2013). *Pakistan Education Statistics 2011-12* (Report No. 252). Islamabad: AEPAM. Retrieved from: http://unesco.org.pk/education/documents/2013/pslm/Pakistan_Education_Statistics.pdf
Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica*, 49, 1417-1426. Retrieved from: http://fmwww.bc.edu/ec-c/S2004/771/NickellEM81.pdf
Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica, 73*, 417-458. doi: 10.1111/j.1468-0262.2005.00584.x
Rockoff, J. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *The American Economic Review, 94*(2), 247-252. doi: 10.1257/0002828041302244
Rothstein, J. (2010). Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement. *Quarterly Journal of Economics, 125*, 175-214. doi: 10.1162/qjec.2010.125.1.175
Sabarwal, S., Evans, D., & Marshak, A. (2013). *The Permanent Textbook Hypothesis: School Inputs and Student Outcomes in Sierra Leone* (Working Paper No. 7021). Washington, DC: The World Bank. Retrieved from: http://documents.worldbank.org/curated/en/806291468299200683/pdf/WPS7021.pdf
Todd, P. E., & Wolpin, K. I. (2003), On the Specification and Estimation of the Production Function for Cognitive Achievement. *The Economic Journal*, 113(485), 3-33. doi:10.1111/1468-0297.00097
UNESCO. (2013). *Status of Teachers in Pakistan* (Report No. 227256). Lahore: UNESCO and ITA. Retrieved from: http://unesco.org.pk/education/documents/2014/publications/Status_of_Teachers_in_Pakistan.pdf

UNESCO. (2014). *Teaching and Learning: Achieving quality for all.* Paris: UNESCO. Retrieved from: http://unesdoc.unesco.org/images/0022/002256/225660e.pdf

Verbeek, M., & Nijman, T. (1992). Testing for Selectivity Bias in Panel Data Models. *International Economic Review, 33*, 681-703. doi:10.2307/2527133

Westbrook, J., Shah, N., Durrani, N., Tikly, C., Khan, W., & Dunne, M. (2009). Becoming a teacher: Transitions from training to the classroom in the NWFP, Pakistan. *International Journal of Educational Development, 29*, 437-444. doi: 10.1016/j.ijedudev.2008.12.001

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT press.