

DOCUMENT DE TRAVAIL

DT/2015-21

Better Teachers, Better Results? Evidence from Rural Pakistan

De Talancé Marine

Better Teachers, Better Results?

Evidence from rural Pakistan

Marine de Talancé*

November 10, 2015

Abstract

Using a gain model with three different levels of fixed effects, this paper empirically estimates the impacts of teachers on students' achievement in three districts in the rural province of Punjab in Pakistan. The model-based results suggest that teachers' factors do explain students' achievement. Increasing teachers' wages could improve schooling quality along with recruiting local teachers with non-permanent contracts. Recruiting local teachers has an important positive impact on students' achievement especially for girls. It could therefore reduce gender based inequalities of academic achievement. Furthermore, our analysis suggests that policy reforms concerning training programs and re-thinking wage policies should be on the agenda of future research.

Keywords: Education, Pakistan, Primary Schooling, Punjab, Quality, Skills, Teachers, Temporary contract

Résumé

Ce papier évalue l'impact des professeurs sur les résultats scolaires des élèves dans trois districts de Punjab au Pakistan. Les résultats de cette étude indiquent que les différences d'enseignants expliquent partiellement la réussite scolaire des élèves. Les élèves qui ont un professeur mieux payé, engagé avec un contrat temporaire et originaire de la même région que l'école ont de meilleurs résultats. Recruter des professeurs locaux est associé à un gain de connaissances supérieur pour les filles par rapport aux garçons. De plus, notre analyse suggère que des réformes concernant les programmes de formation des professeurs devraient être mises en place.

Mots Clés: Qualité de l'éducation, Pakistan, Enseignement primaire, Connaissances cognitives, Professeurs, Contrat temporaire

JEL classification: I21, I25, I28

* (1) PSL, Université Paris-Dauphine, LEDa, DIAL UMR 225, F-75016 Paris, France, (2) IRD, LEDa, DIAL UMR 225, F-75010 Paris, France; e-mail: marine.talance@gmail.com. I would like to thank my supervisors Philippe de Vreyer (DIAL) and Najat El-Mekkaoui (DIAL) for their precious advice and also Florence Arestoff (DIAL) for her very valuable comments.

1 Introduction

There is widespread evidence of the low quality of primary schooling in developing countries where many pupils leave primary schools without basic mathematical and reading skills (Muralidharan and Sundararaman, 2013; NCERT, 2011; WorldBank, 2013).¹ The low quality of schooling in developing countries has substantial economic consequences. Indeed, if the quality of the education delivered is low, students could be incited to drop out of school sooner (Hanushek et al., 2006) leading to a lower human capital accumulation. Future individual earnings would be reduced and inequalities would increase (Behrman et al., 2008; Boissiere et al., 1985; Card and Krueger, 1992; Green and Craig Riddell, 2003; Hanushek, 2005; Murnane et al., 2000). From a macroeconomic point of view, this would translate into lower economic growth (Altinok and Murseli, 2007; Hanushek and Kimko, 2000; Hanushek and Woessmann, 2012). Therefore, identifying the contributors to the quality of schooling is of paramount importance in order to improve the human capital accumulation process in developing countries.

A large literature has tried to assess the determinants of quality of education using education production functions (See Hanushek (2003), Glewwe and Kremer (2006) and Glewwe et al. (2011) for reviews of the existing literature). Whereas research in developed countries has shown that teachers have considerable impacts on students' achievement (Behrman et al., 2008; Card and Krueger, 1992; Chetty et al., 2011; Hanushek, 2011; Nye et al., 2004), most of the studies in developing countries have neglected the impacts of teachers upon children's performance.

The limited progress toward understanding the impacts of teachers upon academic achievement in developing countries stems mainly from data limitations. Because of a lack of panel data with information on students, schools, teachers and households, most research has failed to estimate the effects of teachers on the dynamic process of knowledge acquisition.

In this paper, we determine to what extent students' achievement can be attributed to teachers. We make use of a unique panel dataset on three-, fourth- and fifth-grade students in three districts of Punjab province in Pakistan. Over the course of three years, these children, attending both private and primary schools, were tested in three different subjects (Mathematics, Urdu and English). This project also gathered rich information on households, schools and teachers. We are therefore able to determine to what extent differences in academic achievement between children can be attributed to differences in observable teacher characteristics.

We use the panel dimension of the database to develop an appropriate identification strategy that has three main key features that distinguish it from prior studies in developing countries. First, we are able to develop 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 students' unobserved heterogeneity and particularly for unobserved students' innate abilities, we also make use of students fixed effects comparing students who have been assigned to teachers with different characteristics overtime.

The model-based results point out that teachers' factors have a bigger effect on academic achieve-

ment than background and classroom characteristics. Increasing and reforming teachers' wages is associated with improvement in schooling quality along with recruiting local teachers with non-permanent contracts. Recruiting local teachers could reduce gender based inequalities of academic achievement as it has a larger impact on girls' academic results. Our analysis also suggests that policy reforms concerning training programs should be implemented. These findings are consistent across different specifications, score measurements and sample restrictions.

The remainder of this paper is structured as follow. The second section of this study reviews the relevant literature about teachers and students' achievement. The third section briefly describes the educational context in Pakistan. Section 4 discusses the empirical methodology. In section 5, we describe the database and the variables used in the empirical analysis. Results are presented in section along with various robustness checks. Finally, the last section concludes with the implications for educational policies in Pakistan and for further research.

2 Related literature

Since the influential Coleman report (Coleman et al., 1966), international evidence has shown that traditional inputs-based policies have failed to improve the quality of education (see Hanushek (1986, 2003); Glewwe and Kremer (2006); Murnane and Ganimian (2014) for reviews of the literature). In the absence of natural or randomized experiments, various papers use education production functions. However, until recently, most of these functions did not really take into account the effects of teachers upon child performances.

Recent studies have included teachers' fixed effects within education production functions. Hanushek and Rivkin (2010b) review 10 recent studies in the United States and show that on average an increase of one standard deviation in teacher effectiveness raises students' reading and mathematics scores by respectively 0.13 and 0.17 standard deviation. Other researchers consider the residuals of traditional education production functions as a proxy for teacher quality (Ballou et al., 2004). These results are confirmed by papers showing that teaching effectiveness improves future students' earnings (Behrman et al., 2008; Card and Krueger, 1992; Chetty et al., 2011; Hanushek, 2011). While these approaches are useful to convince policymakers that teachers matter, they do not however point out the determinants of teachers' effectiveness and the policies that should be implemented to improve the quality of schooling.

When specific teacher characteristics are included within education production functions, the results are not convincing. This has been demonstrated in developing countries by Hanushek (2003) and Glewwe et al. (2011). The former reviews 96 education production function estimates implemented before 1990 while the latter provides a review focusing on 79 studies from 1990 to 2010 using randomized controlled trials, difference in differences regressions, regression discontinuity designs or simpler quantitative methods (OLS). These two studies prove that researchers have failed to determine with robustness the characteristics of the teachers influencing students' achievement mainly because of data limitation and econometric issues.

Aslam and Kingdon (2011) use data on 65 schools in Lahore district, Pakistan, in 2002-2003.

They find no evidence proving that observable teachers' characteristics (experience, training, diplomas, absence,...) increase students' achievement. Fehrler et al. (2009) use the SACMEQ and PASEC databases to estimate an education production function in 21 Sub-Saharan countries. They conclude that teachers' educational attainment and professional training have no significant impact on students' achievement because they are not linked with teachers' knowledge. Michaelowa (2001) also uses the PASEC database in 1995-1999 for five African countries and integrates a measure of teachers' job satisfaction within the education production function. She finds that teacher job satisfaction exerts a positive and significant influence on student learning. Aturupane et al. (2013) measure the determinants of academic performance for fourth grade students in Sri Lanka and include very few teachers' characteristics (teacher experience and number of meetings with parents). The effects of teachers' factors disappear when instrumental variables are included.

Therefore, while it is generally acknowledged that teacher quality is a key factor in improving education, to our knowledge, no paper has succeeded in determining convincingly what can explain teachers' effectiveness. Previous studies in developing countries have been hampered by a lack of panel data and therefore they have not been able to take into account the dynamic dimension of the learning process and to address students' selection.

3 Quality of schooling and teachers in Pakistan

3.1 Education in Pakistan

While 10% of primary out-of-school children of the world reside in Pakistan (UNESCO, 2014), many indicators show some 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 primary, secondary and tertiary levels. Nevertheless, Pakistan is still far from achieving the universal educational goal and many Pakistanis remain excluded from the educational system. With a literacy rate of 54.9% and 47.7% of the labor force who is uneducated, Pakistan compares poorly to the average in South Asia. Out of the children attending primary schools, 39% drop out before completing the last grade. The educational system in Pakistan remains discriminating. Low participation in schooling has strong gender, rural and wealth components.²

Even when they go to school, Pakistanis do not necessarily learn the basics. Therefore, even if the universal education goal were achieved, it would not be sufficient if children leaving primary schools are illiterate and innumerate. According to a national survey, by the end of primary school (grade 5), only 54% of the children can read a story in Urdu, 52% are able to read a sentence in English and 47% can divide 3-digit numbers.³ Thus, when they leave school, often even before completing grade 5, most of the students still do not possess the basic knowledge valued by the labor market. This poor quality of schooling may be due to several inefficiencies including poor infrastructures in schools. For instance, the relatively high pupil-teacher ratio⁴ suggests that each teacher has to deal with a large number of children and these pupils receive less attention from their teacher. Physical infrastructures are also often in bad conditions which also can negatively affect the quality of schooling. For instance, 54% of primary schools in Pakistan lack electricity, have unusable latrines and 30% do not have access to clean drinking water (NEMIS-AEPAM, 2011).

3.2 Teachers in Pakistan

Apart from infrastructural deficiencies, the low quality of schooling in Pakistan has also been attributed to a low teaching level (Khamis and Sammons, 2004; Saeed and Mahmood, 2002; Westbrook et al., 2009). Implementing the right policies concerning teachers could improve the quality of primary schooling in Pakistan.

The relatively poor quality of education in Pakistan could partly be attributed to the low level of educational qualifications required to become a primary school teacher. To teach in a primary Pakistani school, teachers need to study at least 10 years (Matric level), a relatively low requirement. Professional preparation of teachers in Pakistan is not standardized and not based on acceptable standards (NEC 2006). Admission to 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). Although the National Education Policy (2009) states that a Bachelor degree in Education (B.Ed.)⁵ should be the minimum required to teach at the elementary level, the old certifications (PTC and CT) remain largely dominant. Indeed, 70% of primary teachers in public schools have followed a PTC or a CT training programs (NEMIS-AEPAM, 2011). Besides this pre-service training, teacher development and continuous training remains voluntary and few teachers benefit from these in-service training programs.

Recruitment of teachers is also problematic as it is based on political pressures and not on merit or on the quality of teachers (Ali, 2000). Before 1997, teachers in Punjab were hired as permanent public servants but this led to politically motivated recruitment and transfers preventing the most competent and qualified teachers from entering the system. In 1997, a ban on hiring new teachers was implemented to face a budgetary crisis and limited financial resources. In 2002, the ban was removed and teachers were hired on a five-year renewable contract. Recruiting contract teachers is a way to deal with the rise of students' enrollment under strong budgetary constraints. These contracts, based on teachers' performance, were implemented to fight against teachers who were not motivated to perform well as their contracts guaranteed employment over their lifetime. This probably affects their motivation to teach and have an impact on students' achievement. They may be better in teaching because they need to show they deserve to have their contracts renewed. However, they may also be less motivated because they feel less considered than permanent teachers.

For decades, the teaching profession has had a low appealing power and social status in Pakistan. Despite some initiatives to improve this status (annual celebration of World Teachers Day, public reward of best teachers, salary packages, improvements in working conditions), teachers in Pakistan still suffer from a compromised social status and this profession is often considered as a semi-profession and is perceived as the last choice of young professionals (UNESCO, 2013; Westbrook et al., 2009). However, teachers in Pakistan are on average as affluent as individuals with at least 10 years of education (Halil et al., 2014). In the public sector, the salaries are governed by the Basic Pay Scales (BPS) system, which is driven by qualification and not by performance. In 2011, a primary

school teacher's monthly salary ranged from Rs. 6,200 (\$ 62.9) to Rs. 17,600 (\$ 178.4). Compared to the public sector, the variations of teachers' salaries in the private sector are tremendous ranging from Rs. 600 (\$ 6.1) per month to Rs. 30,000 (\$ 304.1) (UNESCO, 2013). Contract teachers are hired as 'educators' with possible pay scales ranging from Rs. 4,500 (\$ 45.6) to Rs. 6,500 (\$ 65.9) per month, up to secondary level. Contrary to common opinion, teachers' salaries have been rising recently in South Asia and Pakistan is no exception (UNESCO, 2013).

Andrabi et al. (2011) dedicate a whole chapter to describing teacher quality in Pakistan, using the same database than this paper. They underline the differences between private and public schools. They show that private school teachers have different demographic profiles than public school teachers. They are predominantly young, unmarried female teachers and come from the local area. Their colleagues employed in public institutions are on average better qualified (more educated, trained and experienced) and they are more paid than private school teachers. Differences in wages between public and private school teachers are not only due to differences in education, training and experience. Indeed, Andrabi et al. (2011) show that after controlling for these characteristics, government teachers still earn more than private school teachers. Because of these different profiles, it is crucial to distinguish between government schools and privately managed institutions.

4 Model

In order to determine if differences of learning between children can be attributed to differences in observable characteristics of teachers, an empirical model is derived from a theoretical framework. This section is partially inspired from Todd and Wolpin (2003) and Meghir and Rivkin (2010).

Achievement at a particular age or grade is the result of a cumulative process of skills acquisition. Knowledge acquisition is a production process combining the entire history of individual, family and schooling inputs with a child's initial endowment (innate ability). The achievement of child i at the end of the school year, A_{it} , depends on the history of both family inputs ($H_{i0}, H_{i1}, \dots, H_{it}$) and child characteristics ($C_{i0}, C_{i1}, \dots, C_{it}$), along with past and current school inputs ($QS_{i0}, QS_{i1}, \dots, QS_{it}$) and his innate ability (μ_{i0}).

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

The subscript t on the function a_t allows the impact of the different inputs to depend on the grade of the child. To construct the knowledge accumulation function, we assume that past human capital depreciates at a constant rate $(1 - \delta)$. In other words, between the birth of the child and grade t , the child is assumed to lose every year the same proportion of knowledge. Cognitive skills of child i in grade t , A_{it} , can therefore be represented as the depreciated knowledge in grade $t - 1$ plus the investment made in learning in t , I_{it} .⁶

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

δ refers to the persistence coefficient.

The investment made in grade t is represented by its reduced-form where it is a function of all the inputs in the considered period made by the family (H_{it}), the child himself (C_{it}) and the school (QS_{it}) plus the effect of innate ability (μ_{i0}). Following Todd and Wolpin (2003), endowed ability is assumed to be invariant, meaning that endowment is determined at birth, or at least when the child enters school for the first time, and does not change overtime.

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

In this general model, we allow the effects of inputs to change over time (hence the subscript t on coefficients). Therefore the impacts of the inputs vary with the distance between the current period and the moment when the input was applied (decay rate) and with the grade when it was applied. If the elements of I_{it} are substitutable, equation (2) becomes econometrically :

$$\begin{aligned} A_{it} = & \beta_t H_{it} + \alpha_t C_{it} + \gamma_t QS_{it} + \zeta_t \mu_{i0} + \epsilon_{it} + \\ & \delta[\beta_{t-1} H_{i,t-1} + \alpha_{t-1} C_{i,t-1} + \gamma_{t-1} QS_{i,t-1} + \zeta_{t-1} \mu_{i0} + \epsilon_{i,t-1}] \\ & + \dots + \delta^{t-1}[\beta_1 H_{i,1} + \alpha_1 C_{i,1} + \gamma_1 QS_{i,1} + \zeta_1 \mu_{i0} + \epsilon_{i,1}] + \\ & \delta^t[\beta_0 H_{i,0} + \alpha_0 C_{i,0} + \zeta_0 \mu_{i0} + \epsilon_{i,0}] \end{aligned} \quad (4)$$

$$\Leftrightarrow A_{it} = \sum_{j=0}^t \delta^{t-j} [\beta_j H_{ij} + \alpha_j C_{ij} + \gamma_j QS_{ij} + \zeta_j \mu_{i0} + \epsilon_{ij}] \quad (5)$$

with QS_{i0} a null vector.

Estimating econometrically equation (4) is problematic as genetic endowment (μ_{i0}) is unobserved and no database contains all past and current inputs. Further assumptions are needed in order to reduce the historical data requirements. These assumptions based on the persistence of past inputs on subsequent achievement are presented below moving from the most restrictive to the least restrictive specification.

The simplest specification, known as the contemporaneous model, assumes that only current inputs matter to produce the current child's achievement. It means 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 the history of inputs). This approach has been used by Dee (2004) and Rockoff (2004). In this case, lagged achievement drops out of the knowledge function and equation (4) becomes :

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

where the error term includes the child's endowment ($\epsilon'_{it} = \zeta_t \mu_{i0} + \epsilon_{it}$). In order to estimate equation (6) properly, another assumption is needed : contemporaneous inputs must be uncorrelated to unobserved innate ability which is included in the error term ϵ'_{it} . These two assumptions are unrealistic. First, the skills acquired in previous periods are likely to persist over time (Cunha et al., 2010; Meghir and Rivkin, 2010; Todd and Wolpin, 2003). Furthermore, if parents respond to child's endowment (they invest more in the more "gifted" child), then the second assumption is not valid and equation (6) is inconsistent with OLS. To handle this problem, a fixed-effects model can be used

but this does not solve the question of omitted past inputs variables and the model remains biased (Meghir and Rivkin, 2010). For these reasons, value-added models are generally preferred (Todd and Wolpin, 2003). Two types of value added models can be estimated : the restricted and the unrestricted specifications.

The restricted value added model, also known as the gain score model, assumes that there is no constant decay, or similarly perfect persistence, of past knowledge ($\delta = 1$). Various studies have employed gains in achievement as dependent variables (Ballou et al., 2004; Hanushek et al., 2005; Harris and Sass, 2011; Rivkin et al., 2005).

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

In this model, we consider that the increase (or decrease) of knowledge between $t - 1$ and t is attributable to the investment made in t . For this empirical specification to be valid, inputs have to be uncorrelated with the error term in t and the impacts of each input (and of innate ability) must be independent from the grade when they were applied. Inputs have an immediate impact on achievement that does not depreciate over time. The error terms must also be serially correlated and match the rate of decay of the inputs.⁷

The third model, known as the unrestricted value added-model, make no assumption on the value of the persistence rate δ . This specification can be obtained by subtracting δA_{ig-1} from (4) :

$$A_{ig} = \delta A_{it-1} + \beta H_{it} + \alpha C_{it} + \gamma QS_{it} + \epsilon'''_{it} \quad (8)$$

where the error term includes the child's endowment ($\epsilon'''_{it} = \zeta_t \mu_{i0} + \epsilon_{it}$). Lagged achievement capture the contribution of all previous inputs and past unobservable shocks or endowments. Equation (8) is consistent if the effects of all inputs (including the child's innate ability) are assumed to decline geometrically at the same rate.

Estimating the value-added model is still problematic as there are multiple sources of potential biases. First, prior achievement is correlated with child's endowment by construction which is captured by the individual part of the error term ($\epsilon'''_{it} = e_{it} + \eta_i$). If students with better ability learn faster, $Cov(A_{it-1}, \eta_i) > 0$ and δ will be biased upwards (Andrabi et al., 2011). This bias in the persistence coefficient would lead to bias in input coefficients (Andrabi et al., 2011). As using fixed effects in a dynamic model implies introducing bias, especially if the time dimension is relatively small (Nickell, 1981), the lagged variable needs to be instrumented using the generalized methods of moments developed by Arellano and Bond (1991) or Anderson and Hsiao (1982). Because of data limitation, many of studies have estimated equation 8 by simple OLS ignoring the correlation between lagged achievement and error (Aaronson et al., 2007; Clotfelter et al., 2007; Nye et al., 2004). Andrabi et al. (2011) represent an exception as they deal with this endogeneity issue using appropriate instrumental methods. Following Andrabi et al. (2011), who use the same database than us but look at the effect of private schooling on achievement and not at teacher effects, we estimate the lagged value added model using generalized methods of moments (GMM).

Another issue in estimating equation 8 arises from the omitted variables : ϵ_{it} includes some unmeasured current inputs. Past achievement is likely to be correlated with the error because omitted inputs are subject to choice so optimizing behaviors will create correlation between prior achievement and contemporaneous omitted inputs (ϵ_{it}).

This model is less restrictive than the gain model however due to the low temporal dimension of the database (three waves), we obtain the same results than Andrabi et al. (2011) with both Hansen and Sargan tests rejecting the hypothesis of validity of instruments. We will therefore prefer the gain model (equation 7) and use the unrestricted value added (equation 13) and contemporaneous (equation 6) models as robustness tests.

The vector of school inputs, QS_{it} , contains school-level inputs denoted S_{imt} with m indexing the school; as well as a vector of classroom inputs, P_{it} , and teacher characteristics, T_{ijt} with j indexing the teacher. Classroom inputs refer to peer characteristics and material available in the classroom (books, computers, etc). Teacher characteristics, T_{ijt} , can be decomposed into time-varying teacher characteristics, T_{it}^v (teacher experience, contract, ...) and time-invariant teacher characteristics, T_{it}^i (teacher initial education, gender, ...). The models presented before can be estimated using different fixed effects levels. The value-added model without fixed effects yields :

$$A_{ig} = \delta A_{it-1} + \beta H_{it} + \alpha C_{it} + \gamma T_{ijt} + \phi S_{imt} + \eta P_{it} + \epsilon_{it}''' \quad (9)$$

where ϵ_{it}''' is an error term containing both individual error and idiosyncratic error : $\epsilon_{it}''' = e_{it} + \eta_i$. Assuming that school-level inputs are constant over the time span of the survey, they can be captured using fixed effects.⁸ This approach reduces bias associated with students and teachers sorting into schools (Harris and Sass, 2011). Equation 9 becomes :

$$A_{it} = \delta A_{it-1} + \beta H_{it} + \alpha C_{it} + \gamma T_{it} + \eta P_{it} + s_m + \epsilon_{it}''' \quad (10)$$

where s_m represents school fixed effects.

To control for individual heterogeneity, especially innate ability, it is also possible to include children fixed effects. As underlined before, individual fixed effects are not recommended in an unrestricted value-added model. However, we can add student fixed effects in the gain model (or in the contemporaneous model) yielding :

$$A_{it} - A_{i,t-1} = \beta H_{it}^v + \alpha C_{it}^v + \gamma T_{it} + \eta P_{it} + s_m + c_i + e_{it} \quad (11)$$

where H_{it}^v and C_{it}^v refer to time varying child and family characteristics (age, wealth,...). Invariant family and individual inputs (parental education, child gender, ...) are captured by the individual fixed effects, c_i .

Teacher characteristics can also be estimated through fixed effects yielding :

$$A_{it} - A_{i,t-1} = \beta H_{it}^v + \alpha C_{it}^v + \gamma^v T_{it}^v + \eta P_{it} + \tau_j + s_m + c_i + \epsilon_{it}''' \quad (12)$$

where τ_j refers to teacher fixed effects and captures time-invariant teachers characteristics (T_{it}^i). If this approach controls for potential omitted time-invariant teacher characteristics, it does not allow to identify the impacts of teacher inputs that does not vary overtime. This method is also computationally burdensome. We will nevertheless present briefly the findings obtained using teacher fixed effects.

To sum up, our baseline model is the gain model which is estimated with three different levels of fixed effects. The contemporaneous model and the unrestricted value added specification are used as robustness checks.

5 Database and variables

5.1 Description of the database

To assess the impacts of teachers on primary education, we make use of a rich panel data set coming from the Learning and Educational Achievement in Pakistan Schools (LEAPS) project implemented by Andrabi (Pomona College), Das (World Bank, DEC), Khwaja (Harvard University), Viswanath (World Bank, South Asia) and Zajonc (Harvard University). Between 2001 and 2005, the LEAPS project gathered data on the distribution of schools in rural Punjab but also on the quality of education. Children were tested in three subjects: Mathematics, Urdu (the vernacular) and English. The tests are relatively long (over 40 questions per subject) and were designed in order to cover a broad range of abilities. A small part of questions changes over time but the content covered remains consistent. Annex A describes the tests used by the LEAPS team. Pupils were tested during winter at roughly one year intervals. As the school year ends in the early spring in Pakistan, the test scores gains between grade t and $t - 1$ can be largely attributable to the grade t .

The sample covers 823 schools (first round) in 112 villages in three districts : Attock (North), Faisalabad (Central) and Rahim Yar Khan (South). The sample is not representative of Punjab because 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. All private and public schools within each village boundaries and within a short walk of any village household (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. Multiple questionnaires were distributed to different groups to obtain a complete picture of the educational environment. School principals, teachers and children were surveyed by the LEAPS team.

During the first round, 13,735 children in grade three were tested. They were tracked and retested in grades four (2005) and five (2006). In 2006, new children in grade three were also tested. Between the waves, children could have dropped-out, remained in the same schools and have been promoted or not, switched schools within the village and have been promoted or not, switched to schools outside the village or they could simply have left the village. 95% of the children were however successfully

tracked down. 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. Although the attrition rate is relatively low, we will nevertheless address this issue in the next subsection. LEAPS data confirms that learning is relatively low in Pakistan with only 50% of the children in grade 3 who have mastered the Mathematics curriculum for grade 1.

For the purpose of this study, the original sample of students was reduced to 15,470 children (or 33,685 observations) after keeping schools with at least two different teachers surveyed and after dropping children who were not promoted or double promoted. Among those students, 44% were surveyed three times, 26% were surveyed twice in a row, 3% were surveyed in round one and three but not two and finally 27% were surveyed only once. As using student fixed effects relies on students having different teachers from year to year, estimates of equation 11 could be driven by few observations if most students stay with the same teacher as they progress in grades. However, in our case, 70% of the children surveyed at least twice have changed teachers at least once during the span of the survey.

Across villages, differences between children’s knowledge are significant with some villages performing better on average than others in all subjects (Figure 1). However, when, for each village and for each year, the average scores of the “worst” and the “best” schools of the village are plotted (figure 2), it becomes clear that all villages have both “good” and “bad” schools. Therefore we observe larger differences in knowledge acquisition between schools and not between villages. To assess how much of the difference in test scores can be explained by village attributes, a simple version of variance decomposition is implemented by separately regressing test scores on district, village and school dummies. The residual variation is assumed to be driven by differences across children and measurement error. The R^2 obtained from these regressions are presented in table 1. A large part of scores variation occurs across schools suggesting that the main variables impacting learning processes are at the school level and may be due to teacher characteristics. These estimates also confirm that there are no good or bad villages as the portion of variation attributable to villages is relatively low (between 8% and 12%). The largest differences in learning therefore occur across children enrolled in different schools and not across children living in different villages.

Table 1: Variance decomposition

	Dep. Variable		
	English Test Score	Math Test Score	Urdu Test Score
<i>% of variance explained by</i>			
District Dummies	3.2%	2.9%	1.6%
Mauza Dummies	11.5%	9.5 %	7.5%
School Dummies	42.9%	28.8 %	27.2 %

Source : Author using the first wave of the LEAPS project

Figure 1: Average scores by villages

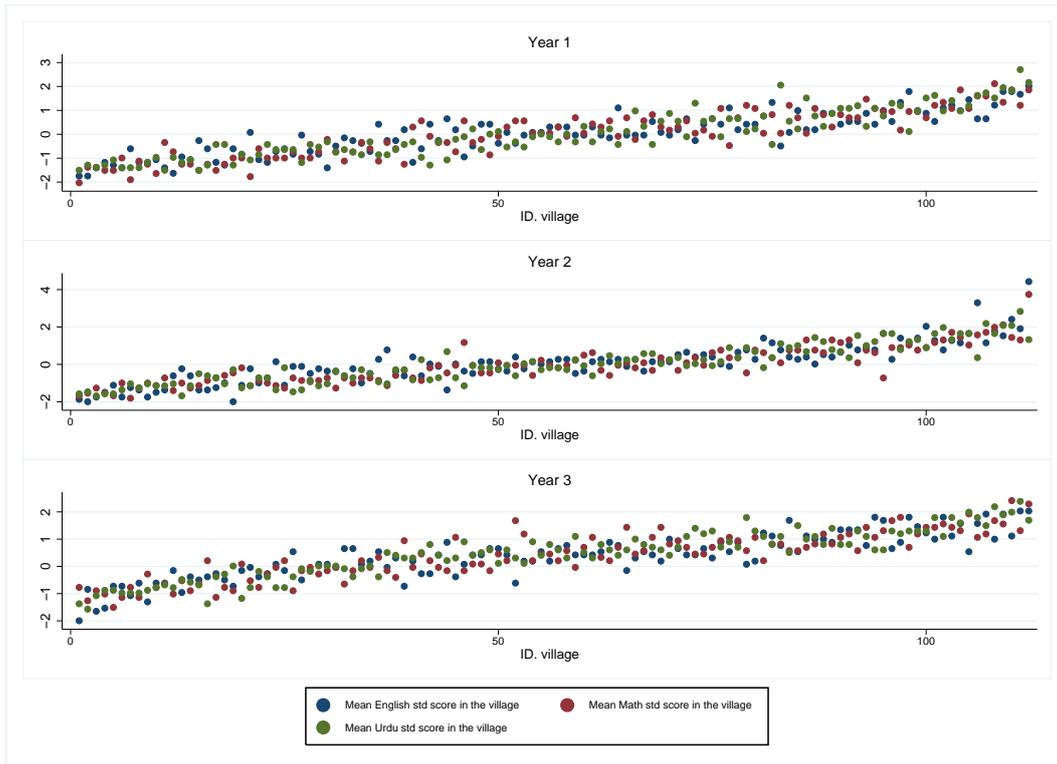
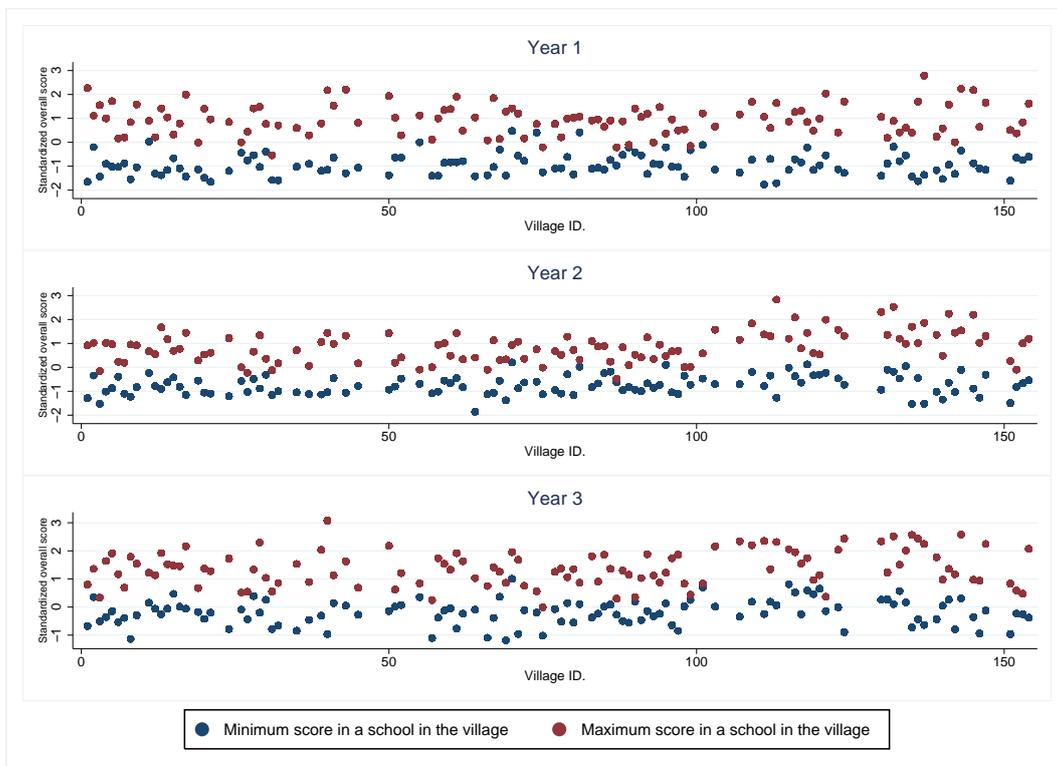


Figure 2: Average scores of the school with the minimum score and the school with the maximum score by villages



Note : In this graph, in each village and for each year, average scores of the school with the average minimum score and the school with the average maximum score are plotted.

5.2 Econometric issues

In order to estimate properly the model specified in equation 7 using the LEAPS panel database, we must address three main empirical challenges : attrition , selection, and assumptions of the gain model. These issues are analyzed in details in Annex B.

5.2.1 Attrition

In Annex B.1, we explore issues related to student and teacher attrition. Given that student attrition is relatively low and that tests suggest that it is not correlated with idiosyncratic errors, estimates using school and student fixed effects are unlikely to be biased by students attrition. We nevertheless provide results for the gain model using the balanced sample as a robustness check. Concerning teacher attrition, evidence in Annex B.1 tends to show that it could bias our estimates. To mitigate this bias, we will include in some regressions teacher fixed effects and provide, as a robustness check, results from the estimates using the balanced sample with only the teachers surveyed during all the waves.

5.2.2 Selection and endogeneity

Besides student and teacher attrition, if students, school resources and teachers are not assigned to schools and classrooms randomly, the specifications presented before could be biased by endogeneity (Ishii and Rivkin, 2009). Non-random assignment consists of four main components : sorting of students to schools, sorting of students to teachers within schools, sorting of teachers to schools and sorting of teachers within contracts. Evidence provided in Annex B.2 shows that children sorting to schools is not likely to bias our estimates, neither is teacher selection of schools. To be sure to alleviate this bias, school fixed estimates are included in all regressions. Teacher selection into contract is not a source of major bias as it depends on variables we control for. Systematic matching of students to teachers is not a major source of bias and the inclusion of student fixed effects alleviates this bias. The only source of potential bias comes from student dynamical matching to teachers.

5.2.3 Assumptions for the value-added model

In the previous section, we presented the different assumptions needed for the gain model to be valid. In Annex B.3, we test whether the effects of the different inputs are constant overtime and whether child's past achievement impacts current inputs. The results tend to validate the assumption according to which parents and schools do not respond to students' past achievement.

5.3 Variables

5.3.1 Dependent Variables

Scores are computed using the Item Reponse Theory method. Contrary to the Classical Test Theory (CTT), it does not assume that each item of a test is equally difficult. Two students who answer the same number of items will not be scored identically unless they have answered the same set of items correctly. In Item Response Theory, the probability of answering correctly to an item is a mathematical function of both individual and item parameters. The estimated scores take into account not only the number of questions answered correctly but also the types of question answered.

Two students will have different scores if one gets a right answer to a more difficult question. IRT approach gives different weights to correct answers depending on the difficulty of the question. These scores are standardized by year and subject. Annex C provides more details on IRT methodology.

5.3.2 Independent Variables

The descriptive statistics of the main explanatory variables are presented in table 2. The first column present the descriptive statistics for the panel, meaning only the children tested three times whereas the three last columns present the statistics for each year. The vector of child characteristics includes gender, age and health. Children in the panel sample are around 10 years old and 46% are girls. The health is measured using the World Health Organization Reference 2007. This index measures body mass in relation with age for children between five and nineteen years old. Children with z-scores below minus two standard deviations from the reference population are considered thin or malnourished, whereas those with z-scores above two standard deviations from the reference are considered overweight or obese. Fourteen percent of the children of the panel sample are underweight and two percent are overweight.

Table 2: Descriptive statistics

	Panel sample		First wave		Second wave		Third wave	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
<i>Child characteristics</i>								
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 : BMI-for-age >2sd	0.02	0.14	0.01	0.12	0.02	0.13	0.03	0.17
Child underweight : BMI-for-age <-2sd	0.14	0.34	0.12	0.32	0.13	0.34	0.14	0.35
<i>Household characteristics</i>								
Dad uneducated	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 more high sec	0.07	0.25	0.06	0.25	0.07	0.26	0.08	0.26
Mum uneducated	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 more high sec	0.01	0.11	0.01	0.12	0.01	0.12	0.01	0.12
Elder siblings (living in the hh)	2.36	2.07	2.36	2.08	2.30	2.04	2.30	2.09
Wealth index	0.03	1.49	-0.33	1.39	-0.10	1.45	0.32	1.53
<i>School characteristics</i>								
Private school	0.29	0.46	0.30	0.46	0.26	0.44	0.29	0.45
<i>Teacher characteristics</i>								
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
Teacher training	0.76	0.42	0.78	0.42	0.76	0.43	0.78	0.41
Continued on next page								

Following the previous table

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)	4854.52	2846.58	4165.63	2274.03	4807.76	2674.63	5696.61	3194.28
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 (days last month)	1.95	2.70	2.15	3.20	1.83	2.60	1.93	2.21
<i>Classroom characteristics</i>								
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 Math 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
<i>District characteristics</i>								
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	
Nb of children	6855		11553		11765		10367	

Note : Panel sample gathers only the children tested during the three waves.

Source : Author, using LEAPS database

The educational history of the parents is also included. Father and mother education backgrounds are included separately as their influence may be different. For instance, one parent may be more involved in educating the child. In the sample, 33% of the children in the sample have an uneducated father and 64% have an uneducated mother. Only 1% have a mother and 7% have a father who had been to more than high secondary school. In order to construct a wealth index, we use household asset indicators to implement a Principal Component Analysis following the approach developed by Filmer and Pritchett (2001). 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.⁹

Various indicators of physical conditions in the classroom are included in our analysis (class size, availability of textbooks, desks, chairs and blackboards). Students in large classes may have less interactions with the teacher and consequently perform poorly. On average, in our sample, 28.6 pupils share the same class which is below the benchmark figure specified by the government. However this average hides a different picture. Indeed, student-teacher ratio is probably a problem in schools in the 90th percentile where it amounts to 50 or more students per teacher. The educational materials

in the class can improve the learning process of the students. However it depends on the quality of the materials and their use. Students are reasonably well equipped in terms of textbooks with 87% of the students within one classroom having a textbook in Urdu, Mathematics and English. Even if most of the pupils have blackboards (84%), many still lack desks (46%) and chairs (81%). We also include a peer variable, the percentage of girls in the class. On average, a child has 46% of girls in his class. This gender composition of the class is highly dependent on the gender of the child as single-sex schools in Pakistan are common. On average, a boy and a girl have respectively 12% and 84% of girls in his or her class.

Concerning the characteristics of teachers which can affect students' achievement, demographic variables are included : gender, gender compared with student's gender and his geographical background. Most of the children surveyed (56%) have a female teacher and 83% have a teacher with the same gender as them. In our panel sample, 62% of the teachers are employed in the same village where they were born.

Apart from demographic indicators, a large number of teachers' characteristics are considered in our model including their education, their training, their experience, the types of their contract, their wages, bonuses and their secondary jobs. On average, teachers have spent 11.5 years studying and 38% have obtained the matric but no higher diplomas. In the panel sample, 44% of the students have a teacher with a PTC certification, 19% with a CT certification, 24% with no training and only 14% with a B.Ed. certification. Teaching experience may impact positively the effectiveness of teaching. However teachers can also exert more effort and be more productive when they start working in a new school to prove they deserve their jobs (Jackson, 2013). On average, teachers have been teaching for 11 years.

As previously underlined, new teachers have been hired with non permanent contracts but the efficiency of this recruitment policy remains uncertain. In our sample, 62% of the sample teachers have permanent contracts. We include the monthly wage of teacher as an explanatory variable. On average, teachers in our sample earn 4,854 Rs (\$ 46.2) per month but we observe a heterogeneity between teachers. In addition to their regular wage, some teachers can receive prizes or bonuses for various reasons (attendance, extra responsibility, pupils' performance, etc.) which can affect their motivation and willingness to perform better. In the sample of teachers, 32% teachers can receive a bonus or a prize but only 9% did receive one.

Teachers working outside the schools may be less involved in their teaching job. In our sample, 15% of the students have a teacher who earn additional income by working in agriculture, 13% have a teacher teaching outside the school, 2% have a teacher running a business and 3% have a teacher with another job. The low attendance of teachers is a crucial issue in developing countries. According to our data, teachers were absent on average two days during the previous month.¹⁰

6 Empirical results

6.1 Students' achievement

Before including teacher observable characteristics into regressions, we regress gain scores on students, schools and teacher fixed effects. These three-way fixed effects estimates give an idea about teacher effects on students' achievement. Table 3 presents the F-statistics of joint significance on teacher fixed effects and shows that teachers are predictors of gains in achievement in each of the three subjects (P-value < 0.01). This table also presents raw standard deviation of teacher fixed effects. This raw standard deviation overestimates teacher effects because of sampling error (Rockoff, 2004; Aaronson et al., 2007). Following Aaronson et al. (2007), we consider that the estimated teacher fixed effect, $\hat{\tau}_j$, has two components the true teacher fixed effect, τ_j , and an error term, ϵ_j , due to sampling : $\hat{\tau}_j = \tau_j + \epsilon_j$. The variance of the estimated teacher effects, $\hat{\sigma}_j^2$, is therefore the sum of the true variance of teacher effect and the average sampling variance : $\hat{\sigma}_j^2 = \sigma_j^2 + N^{-1}\epsilon'\epsilon$. We adjust the variance of teacher fixed effects by subtracting the average sampling variance estimated as the mean of the square of the standard errors of estimated teacher fixed effects. This adjusted standard deviation of teacher effects is lower than the raw measure but it nevertheless shows that moving one standard deviation up on the distribution of teacher fixed effects, increases gains in scores by around 1.

Table 3: Regression with Teacher Fixed Effects

	Dep var : gain scores in		
	English	Math	Urdu
<i>Tests of teacher fixed effects</i>			
F-statistics of joint significance of teacher fixed effects	F(277,8093)= 245.20	F(275,8093)= 3513.12	F(274,8093)= 8048
P-values : Prob>F	0.00	0.00	0.00
<i>Variations in teacher fixed effects</i>			
Raw standard deviation	1.037	1.141	1.081
Adjusted standard deviation	0.939	1.037	0.952
Nb of teacher fixed effects estimates	339	344	346
Nb of observations	12296	12296	12296
Adjusted R ²	0.148	0.167	0.147
FE students	Yes	Yes	Yes
FE schools	Yes	Yes	Yes
Student time varying covariates	Yes	Yes	Yes

Source : Author using the three waves of the LEAPS project

Table 4 presents the increase in the adjusted R² when teacher fixed effects are added to regressions that contain only school and student fixed effects. Across subjects, the variance in score gains accounted for by teacher fixed effects ranges from 17% to 19% suggesting that teachers do explain variance in score gains.

Table 4: Variance Decomposition : variance accounted for by teacher fixed effects

	Dep var : gain scores in		
	English	Mathematics	Urdu
Adjusted R ²			
School and student FE	0.012	0.016	0.013
School, student and teacher FE	0.186	0.191	0.205
Increase in adjusted R ²	0.174	0.175	0.192

Source : Author using the three waves of the LEAPS project

Tables 5 and 6 present the estimates of the gain model (equation (7)) respectively without and with student fixed-effects. Most of the results are robust from one specification to the other. We observe no significant differences in achievement between boys and girls (except in English where girls perform better than boys but this effect is significant only at 10%). Not surprisingly, older children coming from wealthy households have significant higher achievement in all subjects. Parental education has no or low impact on students' achievement but it is probably partly captured by the wealth of the household as parental education and wealth are correlated.

Table 5: Gain model with school fixed-effects

Dep var : IRT gain score - ML	(1) English	(2) Math	(3) Urdu
Girl	0.037* (0.021)	0.011 (0.022)	0.031 (0.023)
Child age	0.020*** (0.006)	0.015*** (0.006)	0.012** (0.006)
Child underweight : BMI-for-age <-2sd	0.001 (0.022)	0.032 (0.023)	0.023 (0.023)
Child overweight : BMI-for-age >2sd	0.029 (0.048)	-0.070 (0.050)	-0.032 (0.051)
Dad less primary	0.009 (0.028)	0.030 (0.030)	-0.008 (0.032)
Dad primary to high sec	0.012 (0.015)	0.016 (0.015)	-0.007 (0.015)
Dad more high sec	0.022 (0.028)	0.080*** (0.027)	-0.013 (0.027)
Mum less primary	-0.016 (0.025)	0.025 (0.025)	-0.024 (0.027)
Mum primary to high sec	-0.027* (0.016)	0.010 (0.016)	-0.004 (0.016)
Mum more high sec	0.026 (0.042)	-0.060 (0.047)	-0.017 (0.046)
Wealth index	0.009* (0.005)	0.019*** (0.006)	0.018*** (0.006)
Female Teacher	-0.092 (0.080)	-0.206** (0.081)	-0.193** (0.086)
Same gender Teacher	-0.005 (0.022)	-0.009 (0.023)	-0.028 (0.024)
Local teacher	0.121*** (0.045)	0.142*** (0.047)	0.107** (0.046)
Teacher Exp	0.012 (0.011)	0.002 (0.010)	0.018* (0.011)
Teacher Exp ²	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)
Teacher education	0.259** (0.130)	0.197 (0.143)	0.204 (0.131)
Teacher education ²	-0.009* (0.005)	-0.007 (0.006)	-0.008 (0.005)
Non-permanent contract	0.250*** (0.072)	0.212** (0.085)	0.252*** (0.077)
PTC training	0.067 (0.057)	-0.148** (0.067)	-0.076 (0.062)
CT training	-0.124* (0.064)	-0.104 (0.072)	-0.141* (0.073)
No training	-0.043 (0.079)	-0.129 (0.095)	-0.049 (0.081)
Log teacher monthly wage	0.258*** (0.064)	0.230*** (0.059)	0.235*** (0.060)
Bonus for pupils' performance	0.062 (0.050)	-0.038 (0.051)	0.070 (0.048)

Continued on next page

Following the previous table

Bonus for other reasons	0.242** (0.108)	-0.053 (0.118)	0.085 (0.106)
Teacher absence	-0.000 (0.007)	-0.005 (0.006)	-0.011* (0.006)
Teacher teaches outside	0.005 (0.051)	0.051 (0.055)	0.022 (0.051)
Teacher other work	0.049 (0.046)	0.014 (0.046)	-0.090** (0.045)
Class size	-0.010*** (0.002)	-0.017*** (0.002)	-0.012*** (0.002)
% girls in the class	0.076 (0.166)	0.067 (0.210)	0.114 (0.233)
% with English books	-0.304*** (0.099)		
% with Math books		-0.310*** (0.111)	
% with Urdu books			-0.421*** (0.102)
% with Desks	0.199*** (0.038)	0.083** (0.039)	0.102*** (0.037)
% with Chairs	0.017 (0.047)	0.023 (0.058)	0.070 (0.046)
% with Blackboards	0.005 (0.046)	-0.031 (0.049)	-0.074 (0.050)
Observations	11181	11181	11181
Adjusted R ²	0.102	0.097	0.078
FE Schools	Yes	Yes	Yes
FE Teachers	No	No	No
FE Students	No	No	No

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

Teacher absence represents self-reported days of absence last month.

Reference categories : Parents have no education, the child is a boy, teacher has followed a BED training program, he cannot receive a bonus, he does not have another job and he has a permanent contract.

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

Table 6: Gain model with school and student fixed-effects

Dep var : IRT gain score - ML	(1) English	(2) Math	(3) Urdu
Child underweight : BMI-for-age <-2sd	0.015 (0.056)	-0.008 (0.058)	0.004 (0.057)
Child overweight : BMI-for-age >2sd	0.027 (0.120)	-0.067 (0.125)	0.063 (0.131)
Wealth index	0.060*** (0.020)	0.089*** (0.021)	0.098*** (0.021)
Female Teacher	-0.107 (0.098)	-0.257*** (0.097)	-0.192* (0.098)
Same gender Teacher	-0.187** (0.090)	-0.063 (0.088)	-0.088 (0.088)
Local teacher	0.110** (0.053)	0.166*** (0.056)	0.064 (0.053)
Teacher Exp	0.006 (0.012)	0.000 (0.011)	0.029** (0.013)
Teacher Exp ²	-0.000 (0.000)	-0.000 (0.000)	-0.001 (0.000)
Teacher education	0.206 (0.160)	0.175 (0.177)	0.172 (0.151)
Teacher education ²	-0.008 (0.007)	-0.007 (0.008)	-0.007 (0.006)
Non-permanent contract	0.269*** (0.084)	0.206** (0.097)	0.269*** (0.087)
PTC training	0.051 (0.068)	-0.155* (0.079)	-0.124* (0.073)
CT training	-0.105 (0.074)	-0.115 (0.084)	-0.206** (0.085)
No training	-0.139 (0.095)	-0.150 (0.116)	-0.162* (0.097)
Log teacher monthly wage	0.262*** (0.079)	0.227*** (0.073)	0.151** (0.073)
Bonus for pupils' performance	0.094 (0.061)	0.015 (0.063)	0.118** (0.059)
Bonus for other reasons	0.258** (0.125)	0.056 (0.137)	0.207* (0.120)
Teacher absence	-0.000 (0.007)	-0.003 (0.007)	-0.013** (0.006)
Teacher teaches outside	0.049 (0.057)	0.078 (0.066)	0.080 (0.058)
Teacher other work	0.044 (0.054)	0.005 (0.056)	-0.118** (0.054)
Class size	-0.008*** (0.003)	-0.015*** (0.002)	-0.010*** (0.002)
% girls in the class	0.437** (0.221)	0.315 (0.287)	0.454 (0.282)
% with English books	-0.250** (0.116)		
% with Math books		-0.271** (0.133)	

Continued on next page

Following the previous table

% with Urdu books			-0.433*** (0.120)
% with Desks	0.157*** (0.044)	0.040 (0.045)	0.054 (0.042)
% with Chairs	0.046 (0.054)	-0.008 (0.070)	0.001 (0.055)
% with Blackboards	0.020 (0.052)	-0.072 (0.056)	-0.063 (0.056)
Observations	11181	11181	11181
Adjusted R ²	0.035	0.049	0.048
FE Schools	Yes	Yes	Yes
FE Teachers	No	No	No
FE Students	Yes	Yes	Yes

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

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

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

Concerning the material available in the classroom, it has a mixed effect on students' achievement. Distributing more textbooks does not seem to be a relevant policy contrary to reducing class size or providing desks. Class sizes negatively impact gains in achievement suggesting that reduced classes could improve learning processes. The mixed and not conclusive effect of textbooks upon students' achievement can be attributed to the lack of clarity of textbooks which makes it hard for the children to understand them and the limited use of the textbook by the teacher (Mohammad and Kumari, 2007). The gender composition of the class, measured by the percentage of girls in the class, positively influences achievement in English probably due to the fact that girls perform slightly better in English than boys on average. This effect probably depend on child time-invariant characteristics such as gender as it disappears when student fixed effects are dropped.

Concerning teachers, women significantly perform poorly compared with male teachers in Mathematics and Urdu. This may be due to a lack of motivation or commitment from female teachers resulting from the fact that most of women declared they will stop teaching after getting married.¹¹ They may also be less involved in their teaching job because they have more responsibilities in their own households.¹² If the benefit from having a same-gender teacher is not obvious, the place of birth of the teacher does play a significant and large role in explaining learning. Teachers who were born in the same village where they teach are more effective, especially in Mathematics. This may be because they speak the same language than the students which facilitates comprehension. They can also be from the same caste and share the same values. Interaction variables were added in order to assess if the positive impact of a local teacher depends on the socioeconomic background of the child but they were not significant. Therefore having of local teacher seems to benefit to a large part of the population whatever their background.

The effects of teacher education and experience are either not significant or relatively low. The positive effect of education on English disappears when individual fixed effects are included. Experience only positively and significantly impacts achievement in Urdu. Interaction variables were added

in order to assess if the impact of education changes according to teacher gender and show that the positive effect of additional education is higher for male teachers (Table 7). On the contrary, the positive impact of experience is higher for female teachers (Table 7). These results must be taken carefully as their magnitudes are relatively low comparatively with other teacher characteristics. Traditional observable teachers characteristics (education and experience) are not associated with substantial improvement in learning. Improving the level of education of teachers in primary schools in Pakistan is probably not the most efficient policy.

Table 7: Gain model - Interactions teacher gender and education/experience

	(1)	(2)	(3)
Dep var : IRT gain score - ML	English	Math	Urdu
<i>Interaction : Gender and education</i>			
Female teacher* Education	-0.079* (0.045)	-0.096** (0.042)	-0.137*** (0.044)
Observations	11181	11181	11181
Adjusted R ²	0.036	0.051	0.051
<i>Interaction : Gender and experience</i>			
Female teacher*Exp	0.006 (0.007)	0.025*** (0.007)	0.016** (0.007)
Observations	11181	11181	11181
Adjusted R ²	0.035	0.052	0.049
FE Schools	Yes	Yes	Yes
FE Teachers	No	No	No
FE Students	Yes	Yes	Yes
<i>Notes</i> : Robust clustered standard errors in parentheses : * $p < .1$, ** $p < .05$, *** $p < .01$			
<i>Reference categories</i> : Teacher has followed a BED training program, he cannot receive a bonus, he does not have another job and he has a permanent contract. <i>Control variables</i> : Same variables than in table 6.			
<i>Source</i> : Author, using the three waves of the LEAPS database.			

Compared with teachers who have followed B.Ed. training programs, teachers with the old PTC and CT certifications are less effective. Training programs seem to be not adapted because they do not implement child-centered pedagogy which is relevant for the context in which teachers work. This is consistent with Osei (2006) who suggests that there is a limited relationship between training and classroom practices in low income countries. Despite the offer of pre-service and in-service training programs, in practice, Pakistani teachers mainly use traditional and conservative teaching methods (Ali, 2000; Mohammed, 2006; Mohammed and Harlech-Jones, 2008; Westbrook et al., 2009). In training programs, theory constitutes 80% of the final training examination whereas teaching practices constitute only 20% of the evaluation of these programs (Westbrook et al., 2009). Trainers themselves often use traditional methods of lecturing and dictating notes therefore discouraging trainees from engaging in debates or asking questions (Ashraf et al., 2005). This translates into poor classroom practices. Many trainers in PTC courses have never taught to children and therefore are not the most qualified to give practical advice (Memon, 2007). Descriptive statistics allow us to compare teaching methods and teacher knowledge between different types of training (Table 8). PTC and CT training programs are clearly inefficient as teachers who have followed them have a lower knowledge

in English, Mathematics and Urdu compared with teachers with a Bachelor in Education. Moreover, they do not seem to have developed adapted pedagogical methods. Reforming training programs is therefore a key feature to improve the quality of primary schooling in Pakistan.

Table 8: Time allocation of teachers and knowledge by training programs

	No training	PTC or CT	BED	Difference	
	Mean	Mean	Mean	Diff (BED-CTorPTC)	T-statistic
English : Teacher Knowledge Score	884.00	883.90	946.93	58.29***	(6.91)
Math : Teacher Knowledge Score	898.28	931.56	955.47	22.05***	(3.62)
Urdu : Teacher Knowledge Score	914.97	925.83	954.11	19.88**	(3.19)
English class time (min daily)	46.51	38.67	54.85	16.18***	(8.04)
Math class time (min daily)	46.14	48.92	51.16	2.24	(1.01)
Urdu class time (min daily)	37.79	43.20	39.27	-3.92	(-1.91)
Islaamiyat class time (min daily)	32.27	34.87	34.23	-0.64	(-0.44)
Science class time (min daily)	30.51	31.37	33.64	2.27	(1.63)
Social class time (min daily)	25.17	25.04	25.78	0.75	(0.61)
Private Tuition (min daily)	37.42	6.13	21.49	15.36***	(6.66)
Religious Activities (min daily)	5.53	7.00	9.03	2.03	(1.55)
Marking Homework/Tests (min daily)	28.36	30.98	26.58	-4.41	(-1.96)
Assembly (min daily)	22.50	22.99	20.68	-2.31*	(-2.30)
Mid break/break/free period (min daily)	26.61	28.94	31.68	2.74	(1.72)
Leisure activities outside of school (min daily)	138.66	104.71	99.51	-5.20	(-0.57)
Community Activities (min daily)	27.09	45.71	38.15	-7.56	(-1.34)
Housework (min daily)	248.54	256.06	242.52	-13.53	(-1.19)
Religious Activities (min daily)	64.61	48.81	86.30	-7.72	(-1.86)
Observations	897	1667	365	2032	

Note : t statistics in parentheses : * $p < .05$, ** $p < .01$, *** $p < .001$.

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

Interestingly, temporary contracts have a positive, significant and strong impact on students' performance. This result, which is consistent with the literature in South Asia (Atherton and Kingdon, 2010; Goyal and Pandey, 2009; Muralidharan and Sundararaman, 2013), may be due to the pressure to have their contracts renewed whereas teachers with permanent contract have no incentives to perform well. However recruiting more teachers with temporary contracts could have a negative impact on students' achievement if these teachers are less trained, educated and experienced and if these negative effects offset the positive effect of pressure. However, when the gain model is estimated without the variables concerning training, education, and experience, the positive effect of temporary

contracts remains significant and strong suggesting that recruiting teachers with temporary contracts could be a way to improve cognitive skills. The impact of having a teacher hired with a temporary contract could depend on how much time he has left before the end of this contract. If the contract arrives to its end soon, it could put pressure on the teacher whereas if the teacher has six years before the end of his contract he would not face such a pressure. To test this, the same gain model is estimated with a dummy indicating whereas the contract expires soon (Table 9). Annex D precises how we create this variable. Our results confirm that the pressure of the end of the contract explains partly why contract teachers perform better at least in Mathematics and Urdu.

Table 9: Gain model - effect of end of the contract

	(1)	(2)	(3)
Dep var : IRT gain score - ML	English	Math	Urdu
Non-permanent contract	0.272*** (0.086)	0.172* (0.099)	0.236*** (0.088)
Non permanent contract*expires soon	-0.010 (0.054)	0.128** (0.065)	0.125** (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 : * $p < .1$, ** $p < .05$, *** $p < .01$
Reference categories : Teacher has followed a BED training program, he cannot receive a bonus, he does not have another job and he has a permanent contract. *Control variables* : Same variables than in table 6
Source : Author, using the three waves of the LEAPS database.

It is plausible that teachers with temporary contract have different teaching methods. To test this assumption, we observe differences in time allocation between contract and permanent teachers (Table 10). Teachers hired as temporary employees spend on average more time teaching English and less time taking breaks or performing community and religious activities. These differences can probably partly explain differences in efficiency between contract and permanent contract teachers.

Table 10: Time allocation of teachers by types of contract

	Permanent contract		Temporary contract		Difference	
	Mean	Sd	Mean	Sd	Diff	T-statistic
English class time (minutes daily)	39.85	33.34	47.45	37.50	-7.604***	(-5.80)
Math class time (minutes daily)	49.36	37.98	46.93	36.96	2.425	(1.73)
Urdu class time (minutes daily)	42.79	34.53	38.64	30.81	4.153***	(3.38)
Islaamiyat class time (minutes daily)	34.31	24.20	33.53	25.39	0.785	(0.85)
Science class time (minutes daily)	31.15	23.58	31.74	27.65	-0.598	(-0.63)

Continued on next page

Following the previous table

Social class time (minutes daily)	24.94	21.25	25.68	21.22	-0.746	(-0.94)
Private Tuition (minutes daily)	4.94	29.26	34.17	70.32	-29.23***	(-15.28)
Religious Activities (minutes daily)	7.00	22.07	6.78	21.86	0.216	(0.26)
Marking Homework/Tests (minutes daily)	29.91	38.70	29.19	40.29	0.713	(0.49)
Assembly (minutes daily)	22.50	17.48	22.46	16.12	0.0437	(0.07)
Mid break/break/free period (minutes daily)	29.44	25.56	26.92	23.94	2.522**	(2.72)
Leisure activities outside of school (minutes daily)	102.67	135.36	128.79	149.73	-26.12***	(-4.23)
Community Activities (minutes daily)	44.42	85.60	33.65	77.15	10.77**	(3.03)
Housework (minutes daily)	263.15	170.28	238.54	169.01	24.60***	(3.34)
Religious Activities (minutes daily)	88.90	63.96	66.93	49.93	21.97***	(8.73)
Observations	1653		1278		2931	

Notes : t statistics in parentheses : * $p < .05$, ** $p < .01$, *** $p < .001$.

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

As contract teachers are less trained and educated on average, we use the results of the teachers at the tests in order to assess differences in knowledge between temporary and permanent contracts. However, descriptive statistics suggest that contract teachers have a broader knowledge in English and a lower level in Mathematics (the difference is not significant for Urdu) (Table 11).

Table 11: Teacher knowledge by types of contract

	Permanent contract		Temporary contract		Difference	
	Mean	Sd	Mean	Sd	Diff	T-statistic
English : Teacher Knowledge Score	885.56	129.82	901.42	97.30	-15.87***	(-3.59)
Math : Teacher Knowledge Score	935.93	101.61	910.41	101.75	25.52***	(6.63)
Urdu : Teacher Knowledge Score	926.92	83.08	925.64	82.01	1.278	(0.41)
Observations	1653		1233		2845	

Notes : t statistics in parentheses : * $p < .05$, ** $p < .01$, *** $p < .001$.

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

Several interaction variables were included into the regressions to test the heterogeneity of the impact of temporary contracts. We observe no difference of the effect of non-permanent contract between local and not local teachers or between male and female teachers.¹³ When temporary contract teacher is interacted with quintiles of wealth of the household, the benefit arising from having a contract teacher is higher in Mathematics for students from the better-off households (Table 12). When we interact education of the teachers and temporary contract dummies, the results show that one additional year of education is more effective for teacher with permanent contracts at least when

it comes to Urdu achievement (Table 12)

Table 12: Gain model - Interaction teacher contract and wealth/teacher education

	(1)	(2)	(3)
Dep var : IRT gain score - ML	English	Math	Urdu
<i>Interaction : temporary contract and household wealth</i>			
Temporary contract*Wealth Quintile 5	0.167 (0.151)	0.402*** (0.154)	-0.026 (0.145)
Temporary contract*Wealth Quintile 4	0.074 (0.122)	0.094 (0.126)	-0.127 (0.119)
Temporary contract*Wealth Quintile 3	0.123 (0.122)	0.220* (0.125)	0.018 (0.120)
Temporary contract*Wealth Quintile 2	0.177 (0.123)	0.242* (0.135)	-0.150 (0.117)
Observations	11181	11181	11181
Adjusted R ²	0.034	0.050	0.046
<i>Interaction : temporary contract and education</i>			
Teacher Education*Temporary Contract	-0.054 (0.041)	-0.015 (0.040)	-0.088** (0.040)
Observations	11181	11181	11181
Adjusted R ²	0.035	0.049	0.049
FE Schools	Yes	Yes	Yes
FE Teachers	No	No	No
FE Students	Yes	Yes	Yes

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

Reference categories : Teacher has followed a BED training program, he cannot receive a bonus, he does not have another job and he has a permanent contract. Control variables : Same variables than in table 6.

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

The more the teacher is paid, the better are his students in all three subjects suggesting that monetary incentives are efficient. However, it seems that this monetary incentive is not as efficient for contract teachers than for permanent teachers (Table 13).

Table 13: Gain model - Interaction teacher contract and wage

	(1)	(2)	(3)
Dep var : IRT gain score - ML	English	Math	Urdu
Non-permanent contract	4.197*** (0.765)	2.991*** (0.757)	1.575** (0.678)
Log wage*Temporary Contract	-0.483*** (0.093)	-0.343*** (0.093)	-0.161* (0.083)
Log teacher monthly wage	0.525*** (0.101)	0.413*** (0.093)	0.239*** (0.087)
Observations	11181	11181	11181
Adjusted R ²	0.042	0.053	0.048

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

Continued on next page

Following the previous table

Reference categories : Teacher has followed a BED training program, he cannot receive a bonus, he does not have another job and he has a permanent contract. *Control variables* : Same variables than in table 6.

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

Students who have a teacher eligible for bonuses because of students' good performance do statistically perform better but only in Urdu and when student fixed effects are included. As shown by Murnane and Ganimian (2014), rewarding teachers on students' performance is efficient only under several conditions. These rewards should be based on test-scores rather than graduation rates and we have no indication if this is the case in our sample. Even if this is the case, the effects of bonuses have been proved to be short lasting as teachers increase their efforts to raise short-run test scores by conducting more preparation sessions but they do not attend the class more nor do they change their pedagogical methods (see Glewwe et al. (2003) for an experiment in Kenya). Being eligible to bonuses for other reasons has significant effects on students' achievement in English.

If the teacher gives private tutoring outside the school, he is not significantly more effective in teaching. If the teacher has another paid job, apart from giving tuition lessons, he tends to be less effective at least in Urdu. These teachers have other potential sources of income and do not rely on their teaching job as much as other teachers. Moreover, they are mainly employed in agriculture, occupations 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 affects significantly and negatively the performance of the children.

The gain model has also been estimated including teacher fixed effects in addition with student and school fixed effects (equation 12). Table 14 present estimates of time-varying teacher observable characteristics using the gain model with three levels of fixed effects. Even after controlling for teacher fixed effects, moving from a teacher with permanent contract to a teacher with a temporary contracts is associated with an increase in gains in Mathematics and Urdu. Teacher wages are still associated with higher achievement gains in all subjects.

Table 14: Gain model with teacher fixed effects

Dep var : IRT gain score - ML	(1) English	(2) Math	(3) Urdu
Child age	0.028 (0.018)	0.039* (0.019)	0.038* (0.018)
Child underweight : BMI-for-age<-2sd	-0.004 (0.057)	-0.003 (0.060)	0.034 (0.058)
Child overweight : BMI-for-age >2sd	-0.094 (0.123)	-0.127 (0.135)	0.004 (0.133)
Wealth index	0.028 (0.020)	0.037 (0.021)	0.059** (0.020)
Same gender Teacher	-0.242** (0.088)	-0.149 (0.088)	-0.125 (0.090)
Teacher Exp	0.012 (0.016)	-0.026 (0.015)	0.044* (0.017)

Continued on next page

Following the previous table

Teacher Exp ²	-0.000 (0.001)	0.001** (0.001)	-0.000 (0.001)
Non-permanent contract	0.071 (0.132)	0.400* (0.170)	0.369** (0.134)
Log teacher monthly wage	0.461*** (0.114)	0.231* (0.099)	0.192* (0.089)
Bonus for pupils' performance	-0.055 (0.089)	0.048 (0.104)	0.029 (0.089)
Bonus for other reasons	0.265 (0.191)	0.432* (0.205)	0.347 (0.186)
Teacher absence	-0.006 (0.009)	0.003 (0.008)	-0.017* (0.007)
Teacher teaches outside	0.105 (0.077)	0.299*** (0.091)	0.076 (0.081)
Teacher other work	0.125 (0.070)	0.073 (0.070)	-0.039 (0.068)
Class size	-0.010** (0.004)	-0.016*** (0.004)	-0.011** (0.003)
% girls in the class	0.385 (0.314)	0.166 (0.439)	-0.058 (0.415)
% with English books	-0.263 (0.158)		
% with Math books		-0.258 (0.163)	
% with Urdu books			-0.380* (0.150)
% with Desks	0.200*** (0.052)	0.142* (0.057)	0.034 (0.052)
% with Chairs	-0.096 (0.077)	-0.136 (0.106)	-0.064 (0.087)
% with Blackboards	0.125* (0.060)	-0.002 (0.063)	0.020 (0.068)
Observations	11719	11719	11719
Adjusted R ²	0.131	0.147	0.134
FE Schools	Yes	Yes	Yes
FE Teachers	Yes	Yes	Yes
FE Students	Yes	Yes	Yes

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

Teacher absence represents self-reported days of absence last month.

Reference categories : Parents have no education, the child is a boy, teacher has followed a BED training program, he cannot receive a bonus, he does not have another job and he has a permanent contract.

To sum up, the main teacher characteristics affecting positively children's achievement are his salary, his geographical background and his contract.

6.2 Robustness checks

6.2.1 Score Measurements

Instead of IRT subject-specific scores computed through maximum likelihood procedures, we use two different measures of scores : the Classical Test Theory measurement and the IRT Expected A Posteriori (EAP) scores. The previous results remain the unchanged (Annex E.1).

6.2.2 Changes in persistence rate

Following Harris and Sass (2011), we examine the robustness of our results to changes in the assumed value of the persistence rate δ (Annex E.2). The positive effects of non-permanent contract teachers (on all three subjects), teacher wages (on all three subjects) and local teacher (on English and Mathematics achievement) remain significant when lower persistence rates are assumed but the magnitude of these impacts vary.

6.2.3 Estimating persistence rates

As some findings depend on the persistence rates assumed, similarly to Andrabi et al. (2011), we estimate the value-added model (equation 8). In order to control for individual heterogeneity that may influence the speed of learning, we use the difference GMM estimator developed by Arellano and Bond (1991) which difference the dynamic lagged value-added model (equation 8) yielding :

$$A_{ig} - A_{it-1} = \delta(A_{it-1} - A_{it-2}) + \beta(H_{it} - H_{it-1}) + \alpha(C_{it} - C_{it-1}) + \gamma(QS_{it} - QS_{it-1}) + (v_{it} - v_{it-1}) \quad (13)$$

where the unobserved fixed effect, η_i is differentiated out. Equation 13 cannot be estimated through simple OLS as A_{it-1} is by construction correlated with v_{it-1} . Following Arellano and Bond (1991), we instrument $A_{it-1} - A_{it-2}$ using two lag periods (A_{i1}, \dots, A_{it-2}) and a certain set of inputs depending on the exogeneity assumption. Results are presented in table 15. Children who changed schools during the span of the survey were excluded and therefore school variables, which are time-invariant are dropped out. The first three columns present the GMM estimates when inputs are considered as exogenous. The lagged achievement in scores is therefore instrumented by scores in $t - 2$ and all current and past inputs. Columns four to six relax the strict exogeneity assumption. Teacher and class inputs are considered as predetermined and the lagged achievement in scores is instrumented by scores in $t - 2$, past and current family and child inputs and past (not current) teacher and classroom inputs. The sample is reduced because only achievement of students in year three, who remain in the same school and have been followed during the three rounds can be used as dependent variables. Like Andrabi et al. (2011), the Hansen test shows that the instrument procedure is weak but we nevertheless use these estimates as robustness tests.

Table 15: Difference in difference GMM Estimates

Dep var : IRT scores - ML	Exogenous Inputs ¹			Predetermined Inputs ¹		
	(1) English	(2) Math	(3) Urdu	(4) English	(5) Math	(6) Urdu

Lagged score in English (IRT-MLE)	0.078** (0.037)			0.125*** (0.038)		
Lagged score in Mathematics (IRT-MLE)		0.211* (0.109)			0.393*** (0.113)	
Lagged score in Urdu (IRT-MLE)			0.311** (0.131)			0.387*** (0.116)
Child age	0.089*** (0.012)	0.099*** (0.015)	0.094*** (0.016)	0.043*** (0.014)	0.032* (0.019)	0.040** (0.018)
Child underweight : BMI-for-age <-2sd	0.047 (0.041)	-0.022 (0.050)	0.008 (0.048)	0.039 (0.045)	-0.027 (0.061)	-0.018 (0.055)
Child overweight : BMI-for-age >2sd	0.216** (0.085)	0.114 (0.102)	0.078 (0.104)	0.191** (0.094)	0.052 (0.124)	0.091 (0.128)
Wealth index	0.055*** (0.015)	0.077*** (0.018)	0.095*** (0.020)	0.037** (0.016)	0.047** (0.021)	0.067*** (0.020)
Female Teacher	0.011 (0.069)	-0.214** (0.083)	-0.073 (0.090)	-0.922*** (0.262)	-1.493*** (0.343)	-1.285*** (0.313)
Same gender Teacher	-0.099 (0.065)	-0.046 (0.077)	-0.112 (0.084)	0.051 (0.225)	0.044 (0.326)	-0.015 (0.284)
Local teacher	0.077** (0.037)	0.173*** (0.045)	0.057 (0.043)	0.075 (0.069)	0.207** (0.091)	0.028 (0.081)
Teacher Exp	-0.002 (0.008)	-0.007 (0.010)	-0.004 (0.010)	-0.038** (0.019)	-0.035 (0.025)	-0.049** (0.023)
Teacher Exp ²	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.001** (0.001)	0.001 (0.001)	0.002** (0.001)
Teacher education	0.095 (0.090)	0.054 (0.124)	0.004 (0.098)	1.107** (0.465)	1.959*** (0.733)	1.239** (0.594)
Teacher education ²	-0.004 (0.004)	-0.001 (0.005)	0.001 (0.004)	-0.046** (0.020)	-0.080*** (0.031)	-0.050** (0.025)
Non-permanent contract	0.078 (0.065)	0.133* (0.079)	0.128 (0.078)	0.257* (0.144)	0.465** (0.204)	0.453** (0.183)
PTC training	-0.023 (0.049)	-0.121** (0.059)	-0.088 (0.057)	0.325*** (0.109)	0.375** (0.154)	-0.060 (0.139)
CT training	-0.112** (0.055)	-0.105 (0.067)	-0.131* (0.071)	0.170 (0.125)	0.393** (0.177)	-0.225 (0.168)
No training	-0.072 (0.067)	-0.041 (0.096)	-0.086 (0.083)	0.242 (0.188)	0.939*** (0.276)	0.222 (0.248)
Log teacher monthly wage	0.195*** (0.048)	0.252*** (0.064)	0.157** (0.064)	0.504*** (0.112)	0.731*** (0.158)	0.525*** (0.146)
Bonus for pupils' performance	0.046 (0.047)	0.010 (0.054)	0.067 (0.053)	0.004 (0.106)	0.404*** (0.148)	0.327** (0.139)
Bonus for other reasons	-0.035 (0.083)	-0.132 (0.105)	-0.051 (0.095)	-0.142 (0.121)	-0.187 (0.170)	-0.028 (0.151)
Teacher absence	-0.001 (0.005)	0.003 (0.006)	-0.007 (0.006)	-0.009 (0.007)	-0.011 (0.009)	-0.019* (0.010)
Teacher teaches outside	-0.016 (0.040)	-0.010 (0.051)	0.005 (0.047)	-0.018 (0.085)	-0.064 (0.110)	-0.135 (0.104)
Teacher other work	0.077* (0.041)	0.058 (0.046)	-0.022 (0.046)	0.065 (0.064)	0.110 (0.079)	0.003 (0.072)

Class size	-0.009*** (0.002)	-0.010*** (0.002)	-0.009*** (0.002)	-0.016*** (0.004)	-0.016*** (0.005)	-0.017*** (0.005)
% girls in the class	0.051 (0.167)	0.048 (0.236)	0.093 (0.282)	-1.134* (0.604)	-1.539* (0.900)	-0.469 (0.745)
% with English books	-0.125 (0.083)			-0.104 (0.125)		
% with Math books		-0.005 (0.109)			0.331** (0.163)	
% with Urdu books			-0.231** (0.104)			-0.313** (0.154)
% with Desks	0.046 (0.032)	-0.014 (0.034)	0.020 (0.034)	0.010 (0.058)	-0.131* (0.073)	-0.082 (0.067)
% with Chairs	-0.012 (0.040)	0.049 (0.050)	-0.008 (0.045)	0.129* (0.067)	0.024 (0.103)	0.127 (0.082)
% with Blackboards	0.019 (0.036)	-0.075* (0.044)	-0.030 (0.044)	-0.149*** (0.057)	-0.247*** (0.079)	-0.060 (0.070)
Observations	2235	2235	2235	2518	2518	2518
Hansen Statistics	94.165	114.157	92.229	58.673	87.823	93.696
Hansen p-value	0.000	0.000	0.000	0.000	0.000	0.000
FE Schools	No	No	No	No	No	No
FE Teachers	No	No	No	No	No	No
FE Students	No	No	No	No	No	No

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

¹ represents GMM estimates when all inputs are considered as exogenous except the lagged scores. Lagged scores are instrumented by scores in $t - 2$ and current and past inputs. ² represents GMM estimates when parental and child inputs are considered as exogenous but lagged scores and teacher and class inputs are considered as endogenous. Lagged scores are instrumented by scores in $t - 2$ and past inputs. Children who changed schools are excluded from this sample *Reference categories* : Teacher has followed a BED training program, he cannot receive a bonus, he does not have another job, he has been absent at least one day during the past month and he has a permanent contract.

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

Consistent with Andrabi et al. (2011), we found relatively low persistence rates ranging from 0.09 to 0.46. Children lose more than half of their achievement in a single year. As underlined by Andrabi et al. (2011), this low persistence of knowledge is unlikely to be driven by mechanical mechanisms such as cheating because the test was implemented by the external LEAPS team and it was low stakes, not part of the standard educational infrastructure preventing teachers from teaching just for the test.

The results from the value-added specification confirm that local teachers are associated with better achievement in Mathematics and that non permanent teachers and better paid tend to be more effective in teaching all subjects.

6.2.4 Gender and types of schools

We also run robustness checks using the gain model (equation 7) to see if the variables impacting students' performance vary according to child gender or to the type of school attended. The previous results remain unchanged (Annex E.3). The benefit from being taught by a local and female teacher is stronger for girls than boys suggesting that hiring local and female teachers could reduce gender gap in academic achievement.

6.2.5 Balanced Samples

In addition with these sub-sample breakdowns, to deal with attrition, estimates for the gain model using the balanced student and teacher panel samples are presented in Annex E.4. The previous results are robust to this sample change.

6.3 Analysis of teachers' wage and teacher recruitment

As previously underlined, raising teachers' wages improves quality of teaching. By analyzing deeply the determinants of teachers' wage it is possible to determine if the criteria used to establish teachers' salaries are relevant with the factors influencing students' performance (Table 16).¹⁴

Table 16: Analysis of teachers' wage

	Dep var : Monthly log of teacher's wage		
	All teachers	Public	Private
Female Teacher	-0.147*** (0.036)	-0.104*** (0.036)	-0.353*** (0.110)
Lagged students' mean English Score (IRT-ML)	0.001 (0.021)	0.016 (0.019)	0.026 (0.069)
Lagged students' mean Math Score (IRT-ML)	0.044 (0.031)	0.024 (0.024)	0.089 (0.118)
Lagged students' mean Urdu Score (IRT-ML)	-0.021 (0.029)	-0.016 (0.024)	-0.115 (0.116)
Nb students	0.001*** (0.000)	0.001 (0.000)	0.003* (0.001)
Diploma : Matric	-0.413*** (0.086)	-0.279*** (0.070)	-1.032*** (0.307)
Diploma : BA-BSc	-0.108* (0.064)	-0.080* (0.043)	-0.552* (0.310)
Diploma : FA-FSc	-0.361*** (0.082)	-0.247*** (0.066)	-0.889*** (0.309)
Teacher Experience	0.045*** (0.010)	0.016 (0.015)	0.042** (0.017)
Teacher Experience ²	-0.001** (0.000)	0.000 (0.000)	-0.001** (0.000)
Local teacher	-0.130*** (0.034)	-0.080** (0.036)	-0.197** (0.086)
PTC training	0.186*** (0.067)	0.349*** (0.102)	0.010 (0.086)
CT training	0.160** (0.073)	0.346*** (0.098)	0.199 (0.175)
BED training	0.229*** (0.077)	0.438*** (0.091)	0.070 (0.198)
Teacher non-permanent contract	-0.369*** (0.080)	-0.527*** (0.104)	-0.007 (0.125)
Teacher absence (days last month)	0.008 (0.005)	0.008*** (0.003)	0.013 (0.022)
District 2	-0.151*** (0.044)	-0.084* (0.046)	-0.247*** (0.080)
District 3	-0.096***	-0.088***	-0.077

	(0.035)	(0.031)	(0.085)
Adjusted R ²	0.855	0.626	0.463
Observations	766	577	189

Notes : Robust clustered (by teachers) standard errors in parentheses : * $p < .1$, ** $p < .05$, *** $p < .01$
Reference categories : Male teacher who has a master or higher diploma, has no training and a permanent contract.
Source : Author, using the three waves of the LEAPS database.

The main factors affecting teachers' wages in both public and private schools are gender and diploma. The negative bias against women is more important in private schools probably because, in private schools, teachers have wages partly based on employment opportunities outside teaching and women have limited employment opportunities in the private sector in Pakistan (Andrabi et al., 2011). The rewards to training are tremendous in public school whereas experience is more determinant in private schools. Teachers with permanent contracts are also more paid in public schools but not in private institutions. In both types of schools, local teachers are less paid and similar to Andrabi et al. (2011), we find that the difference is higher in private schools. Local teachers are significantly less paid maybe because these teachers value safety and convenience of working near where they live along with lower travel costs and therefore they are willing to accept lower salaries.

As suggested by this study, all these criteria determining teachers' wages, are no guarantee of teachers' quality. Wage policy is thus far from being the most efficient. Indeed, teachers' performance, measured by lagged academic results of their previous students, is not rewarded by higher wages. Therefore, teachers receive no financial incentive to increase students' achievement which probably undermines their motivation. Teachers could be systematically evaluated and their wages could be linked with the progresses made by their students. Unless training programs are reformed, there is no reason for the teachers' wages to depend on training.

In the LEAPS survey, school principals have described what are the main characteristics they value when they choose to recruit a teacher (Table 17). Teachers are hired mainly based on their experience, education and training qualifications whereas this survey tends to show that these characteristics are not those improving efficiently learning processes. It is therefore crucial to rethink recruitment policies and to find a way to encourage school principals to hire teachers with the "right" qualities.

Table 17: Descriptive statistics : policies about teachers ' recruitment

	% of head teachers declaring that					
	First criteria is			Second criteria is		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Highly educated teacher	33	29	23	29	23	24
Trained teacher	27	26	24	31	37	33
Highly experienced teacher	21	26	31	18	18	19
Local teacher	7	6	6	8	8	8

Continued on next page

Following the previous table

Female teacher	4	6	5	5	5	6
Young teacher	2	1	1	6	5	5
Teacher from wealthy hh	2	0.4	0.5	2	1	0.5
Teacher enrolls new student	5	6	9	5	4	6

Lecture : During the first wave of the survey, 33% of head teachers declared that the first reason motivating their decision to hire a teacher is his level of education.

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

7 Discussion and conclusion

In this paper, we analyze the characteristics of teachers impacting the skills of students in primary schooling in three districts of the Punjab province in Pakistan, using a gain model. We include both school and student fixed effects in order to deal with non-random sorting of students and teachers and to control for unobserved heterogeneity.

This study provides strong evidence on the relationship between teachers and acquisition of skills. The results suggest that teachers are one of the main drivers of learning. Furthermore, findings show that some observable teacher characteristics are associated with students' achievement. Teachers with temporary contracts seem to perform better than permanent teachers. Locally recruited teachers also tend to be more effective. Finally, our results also suggest that teachers' wage policy could be a tool to motivate teachers and improve the quality of schooling. All these policies are easier to implement than the traditional policies such as raising teacher education or experience.

Evidence on the relationship between teachers' wages and students' achievement raises the question of the design of wages. The current wage policy appears to be far from being the most efficient as it is not linked to the performance of teachers but to characteristics that do not drive teachers' efficiency. It is therefore crucial to reexamine the wage policy in order to use it as an efficient tool to motivate teachers. Future research could use experiments to determine the causal effect of linking teacher performance and teacher wages upon students' achievement.

Experience and education of teachers have a relatively low impact on students' achievement. Nevertheless, it would be interesting to look separately at total teaching experience and experience in the current school. Indeed, it is important to differentiate both as schools probably have different teaching or management methods. When they remain a long time in the school, teachers may learn teaching methods adapted to the specific children in the school (Boyd et al., 2005; Hanushek et al., 1999). If the results confirmed this intuition, retaining teachers should also be one of the priorities of school principals. Further research should therefore assess why teachers leave schools to join another one in order to reduce teachers' attrition and improve quality of education.

Recruiting local teachers is proven to improve the quality of learning especially in Mathematics. It would be interesting to determine the mechanisms driven this relationship. It may be because

teachers come from the same caste than their students, they share the same values or because they speak the regional language. Because of data limitation, unfortunately, we were not able to perform this analysis.

Another pathway to increase the quality of schooling resides in contract teachers and strengthening the pressure they face to have their contracts renewed through systematic and external evaluations. Future research should look into the most efficient ways to assess teachers' quality in order to decide whether or not to renew their contracts. It could also be interesting to determine if permanent teachers who are periodically evaluated by supervisors tend to perform as well as contract teachers. If so, two types of policies could be implemented either recruiting more contract teachers or strengthening supervisions on both permanent and contract teachers.

Finally, reforming teachers training courses is necessary to improve the quality of primary education in Pakistan. It is therefore crucial to implement further analysis in order to pinpoint the deficiencies of current programs and reform them in order to improve teachers' pedagogy.

Notes

¹The quality of schooling in this paper is defined through a value-added approach, meaning that a high quality education is achieved when students show substantial progress in learning between the beginning and the end of the year (Harvey and Green, 1993; Tymms, 1999). To measure these learning outcomes, we use, as it is common in the literature (Boissiere et al., 1985; Behrman et al., 2008; Card and Krueger, 1992; Hanushek and Kimko, 2000; Hanushek and Woessmann, 2012), standardized test scores in different subjects. We are fully aware of the limits of such a measurement. Indeed, test scores do not encompass other possible objectives of the educational system such as the development of specific attitudes and behaviors (flexibility, adaptability, curiosity, empathy,...) or values (solidarity, tolerance, gender equality, respect for difference, human rights and human life). Nevertheless, this measure allows us to quantify the knowledge acquired within the schooling life and for this reason it goes beyond the simplistic traditional conception of human capital as the years of schooling completed.

²Only 66.4% of primary age girls actually attend a primary school. 70.2% of primary school age children living in urban areas are enrolled in a primary school whereas it is the case of only 55.7% of children in rural areas (DHS, 2013). Children coming from poorer households also participate less in schooling. 81.6% of the primary school age children belonging to the 20% richest households attend a primary school whereas this proportion only amounts 35.8% for the children belonging to the 20% poorest households (DHS, 2013).

³To assess the quality of education in whole Pakistan, we use the ASER (Annual Status of Education Report) data. Every year since 2008, this project estimates the skills acquired by 5-16 years old children in 138 rural districts of Pakistan and, since 2011, in urban centers. In each district, 30 villages are selected randomly. In each village, 20 households are randomly selected and the children of these households are tested in Reading (Urdu, Pashto and Sindhi), Mathematics and English.

⁴ On average, in Pakistan, 37 pupils attending primary school share the same teacher (NEMIS-AEPAM, 2011).

⁵B.Ed, Bachelor in Education is a one year post graduate qualification program after a Bachelor degree.

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

⁷The model also holds if there is no decay.

⁸Data 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 the very low time variation of school characteristics, we make use of school fixed effects.

⁹On purpose, we excluded agricultural assets because in our opinion it does not mean households owning them are richer.

¹⁰As the days of absenteeism are self-reported by the teacher, they may be underestimated.

¹¹In our teacher sample, 27% of female teachers (vs. 16% of the male teachers) declare they would stop teaching once they get married and 47% (3%) declare they would stop depending on their spouses' decision.

¹²Female teachers spend on average five hours doing housework per day whereas male teachers spend only three hours doing so.

¹³ The interaction variables were only negatively significant for female teachers at 10% for achievement in Mathematics.

¹⁴The results are robust if we integrate only one subject specific students' mean score.

References

- Aaronson, D., Barrow, L., and Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, 25(1):95–135.
- Ali, M. A. (2000). Supervision for teacher development: an alternative model for Pakistan. *International Journal of Educational Development*, 20(3):177–188.
- Altinok, N. and Murseli, H. (2007). International database on human capital quality. *Economics Letters*, 96(2):237–244.
- Anderson, T. W. and Hsiao, C. (1982). Formulation and estimation of dynamic models using panel data. *Journal of Econometrics*, 18(1):47–82.
- Andrabi, T., Das, J., and Khwaja, A. I. (2008). A dime a day: The possibilities and limits of private schooling in Pakistan. *Comparative Education Review*, 52(3):329–355.
- Andrabi, T., Das, J., Khwaja, A. I., Vishwanath, T., and Zajonc, T. (2007). Learning and educational achievements in Punjab schools (leaps): Insights to inform the education policy debate. *World Bank, Washington, DC*.
- Andrabi, T., Das, J., Khwaja, A. I., and Zajonc, T. (2011). Do value-added estimates add value? accounting for learning dynamics. *American Economic Journal: Applied Economics*, pages 29–54.
- Andrabi, T. R. S., Das, J., and Khwaja, A. I. (2006). *A dime a day: The possibilities and limits of private schooling in Pakistan*, volume 4066. World Bank Publications.
- Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2):277–297.
- Ashraf, D., Shamatov, D. A., Tajik, M. A., Vazir, N., et al. (2005). Reconceptualization of teacher education experiences from the context of a multicultural developing country. *Journal of Transformative Education*, 3(3):271–288.
- Aslam, M. and Kingdon, G. (2011). What can teachers do to raise pupil achievement? *Economics of Education Review*, 30(3):559–574.
- Atherton, P. and Kingdon, G. (2010). The relative effectiveness and costs of contract and regular teachers in India. *Institute of Education, University of London*.
- Aturupane, H., Glewwe, P., and 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(1):2–37.
- Baker, F. B. and Kim, S.-H. (2004). *Item response theory: Parameter estimation techniques*. CRC Press.
- Ballou, D., Sanders, W., and Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, 29(1):37–65.
- Barrera-Osorio, F. and Raju, D. (2014). Teacher performance pay: Experimental evidence from Pakistan. *Washington, DC: The World Bank*.

- Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, 70:9.
- Behrman, J. R., Ross, D., and Sabot, R. (2008). Improving quality versus increasing the quantity of schooling: Estimates of rates of return from rural pakistan. *Journal of Development Economics*, 85(1-2):94–104.
- Betts, J. R., Zau, A., and Rice, L. (2003). *Determinants of student achievement: New evidence from San Diego*. Public Policy Institute of California San Francisco, CA.
- Boissiere, M., Knight, J. B., and Sabot, R. H. (1985). Earnings, schooling, ability, and cognitive skills. *American Economic Review*, 75(5):1016–30.
- Bowles, S., Gintis, H., and Osborne, M. (2001). The determinants of earnings: A behavioral approach. *Journal of Economic Literature*, 39(4):1137–1176.
- Boyd, D., Grossman, P., Lankford, H., Loeb, S., and Wyckoff, J. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. *Education*, 1(2):176–216.
- Boyd, D., Lankford, H., Loeb, S., and Wyckoff, J. (2005). Explaining the short careers of high-achieving teachers in schools with low-performing students. *American economic review*, pages 166–171.
- Card, D. and Krueger, A. B. (1992). Does school quality matter? returns to education and the characteristics of public schools in the united states. *Journal of Political Economy*, 100(1):1–40.
- Chetty, R., Friedman, J. N., and Rockoff, J. E. (2011). The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood. Technical report, National Bureau of Economic Research.
- Clotfelter, C. T., Ladd, H. F., and Vigdor, J. (2005). Who teaches whom? race and the distribution of novice teachers. *Economics of Education review*, 24(4):377–392.
- Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources*, 41(4):778–820.
- Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, 26(6):673–682.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., and York, R. (1966). Equality of educational opportunity. *Washington, D.C.*
- Cunha, F., Heckman, J. J., and Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931.
- Dee, T. S. (2004). Teachers, race, and student achievement in a randomized experiment. *Review of Economics and Statistics*, 86(1):195–210.
- DHS (2013). Pakistan demographic and health survey 2012-2013. Technical report, National Institute of Population Studies (NIPS) and ICF International.

- Fehrler, S., Michaelowa, K., and Wechtler, A. (2009). The Effectiveness of Inputs in Primary Education: Insights from Recent Student Surveys for Sub-Saharan Africa. *Journal of Development Studies*, 45(9):1545–1578.
- Feng, L. (2009). Opportunity wages, classroom characteristics, and teacher mobility. *Southern Economic Journal*, pages 1165–1190.
- Filmer, D. and Pritchett, L. H. (2001). Estimating wealth effects without expenditure data—or tears: An application to educational enrollments in states of india. *Demography*, 38(1):115–132.
- Fitzgerald, J., Gottschalk, P., and Moffitt, R. A. (1998). An analysis of sample attrition in panel data: The michigan panel study of income dynamics.
- Garet, M. S., Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W., Uekawa, K., Falk, A., Bloom, H. S., Doolittle, F., et al. (2008). The impact of two professional development interventions on early reading instruction and achievement. ncee 2008-4030. *National Center for Education Evaluation and Regional Assistance*.
- Garet, M. S., Wayne, A. J., Stancavage, F., Taylor, J., Walters, K., Song, M., Brown, S., Hurlburt, S., Zhu, P., Sepanik, S., et al. (2010). Middle school mathematics professional development impact study: Findings after the first year of implementation. ncee 2010-4009. *National Center for Education Evaluation and Regional Assistance*.
- Glewwe, P., Ilias, N., and Kremer, M. (2003). Teacher incentives. Technical report, National Bureau of Economic Research.
- Glewwe, P. and Kremer, M. (2006). *Schools, Teachers, and Education Outcomes in Developing Countries*, volume 2 of *Handbook of the Economics of Education*, chapter 16, pages 945–1017. Elsevier.
- Glewwe, P. W., Hanushek, E. A., Humpage, S. D., and Ravina, R. (2011). School Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010. NBER Working Papers 17554, National Bureau of Economic Research, Inc.
- Goyal, S. and Pandey, P. (2009). How do government and private schools differ? findings from two large indian states.
- Green, D. A. and Craig Riddell, W. (2003). Literacy and earnings: an investigation of the interaction of cognitive and unobserved skills in earnings generation. *Labour Economics*, 10(2):165–184.
- Halil, D., Beteille, T., Riboud, M., and Deolalikar, A. (2014). Student learning in south asia. challenges, opportunities, and policy priorities. Technical report, World Bank.
- Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of economic literature*, pages 1141–1177.
- Hanushek, E. A. (2003). The failure of input-based schooling policies. *The economic journal*, 113(485):F64–F98.
- Hanushek, E. A. (2005). The economics of school quality. *German Economic Review*, 6(3):269–286.

- Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review*, 30(3):466–479.
- Hanushek, E. A., Kain, J. F., O’Brien, D. M., and Rivkin, S. G. (2005). The market for teacher quality. NBER Working Papers 11154, National Bureau of Economic Research, Inc.
- Hanushek, E. A., Kain, J. F., and Rivkin, S. G. (1999). Do higher salaries buy better teachers? Technical report, National bureau of economic research.
- Hanushek, E. A., Kain, J. F., and Rivkin, S. G. (2004). Why public schools lose teachers. *Journal of human resources*, 39(2):326–354.
- Hanushek, E. A. and Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American economic review*, pages 1184–1208.
- Hanushek, E. A., Lavy, V., and Hitomi, K. (2006). Do students care about school quality? determinants of dropout behavior in developing countries. Technical report, National Bureau of Economic Research.
- Hanushek, E. A. and Rivkin, S. G. (2006). Teacher quality. *Handbook of the Economics of Education*, 2:1051–1078.
- Hanushek, E. A. and Rivkin, S. G. (2010a). Constrained job matching: Does teacher job search harm disadvantaged urban schools? Technical report, National Bureau of Economic Research.
- Hanushek, E. A. and Rivkin, S. G. (2010b). Generalizations about using value-added measures of teacher quality. *The American Economic Review*, 100(2):267–271.
- Hanushek, E. A. and Woessmann, L. (2012). Do better schools lead to more growth? cognitive skills, economic outcomes, and causation. *Journal of Economic Growth*, 17(4):267–321.
- Harris, D. N. and Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of public economics*, 95(7):798–812.
- Harvey, L. and Green, D. (1993). Defining quality. *Assessment & Evaluation in Higher Education*, 18(1):9–34.
- Hornig, E. L. (2009). Teacher tradeoffs: Disentangling teachers’ preferences for working conditions and student demographics. *American Educational Research Journal*, 46(3):690–717.
- Ishii, J. and Rivkin, S. G. (2009). Impediments to the estimation of teacher value added. *Education*, 4(4):520–536.
- Jackson, C. K. (2013). Match quality, worker productivity, and worker mobility: Direct evidence from teachers. *Review of Economics and Statistics*, 95(4):1096–1116.
- Jacob, B. A. and Lefgren, L. (2004). The impact of teacher training on student achievement quasi-experimental evidence from school reform efforts in chicago. *Journal of Human Resources*, 39(1):50–79.
- Jacob, B. A. and Lefgren, L. (2008). Can principals identify effective teachers? evidence on subjective performance evaluation in education. *Journal of Labor Economics*, 26(1):101–136.

- Jamison, E. A., Jamison, D. T., and Hanushek, E. A. (2007). The effects of education quality on income growth and mortality decline. *Economics of Education Review*, 26(6):771–788.
- Jepsen, C. (2005). Teacher characteristics and student achievement: Evidence from teacher surveys. *Journal of Urban Economics*, 57(2):302–319.
- Jones, A. M., Rice, N., and Dias, P. R. (2011). Long-term effects of school quality on health and lifestyle: Evidence from comprehensive schooling reforms in england. *Journal of Human Capital*, 5(3):342–376.
- Kane, T. J., Rockoff, J. E., and Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? evidence from new york city. *Economics of Education Review*, 27(6):615–631.
- Kane, T. J. and Staiger, D. O. (2008). Estimating teacher impacts on student achievement: An experimental evaluation. Technical report, National Bureau of Economic Research.
- Khamis, A. and Sammons, P. (2004). Development of a cadre of teacher educators: some lessons from pakistan. *International Journal of Educational Development*, 24(3):255–268.
- Koedel, C. and Betts, J. R. (2011). Does student sorting invalidate value-added models of teacher effectiveness? an extended analysis of the rothstein critique. *Education*, 6(1):18–42.
- Meghir, C. and Rivkin, S. G. (2010). Econometric methods for research in education. Technical report, National Bureau of Economic Research.
- Memon, G. R. (2007). Education in pakistan: The key issues, problems and the new challenges. *Journal of Management and Social Sciences*, 3(1):47–55.
- Michaelowa, K. (2001). Primary education quality in francophone sub-saharan africa: Determinants of learning achievement and efficiency considerations. *World Development*, 29(10):1699–1716.
- Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy*, 66:281.
- Mohammad, R. and Kumari, R. (2007). Effective use of textbooks: A neglected aspect of education in pakistan. *Journal of Education for International Development*, 3(1):1–12.
- Mohammed, R. F. (2006). Problems of teachers’ re-entry in schools after in-service education.
- Mohammed, R. F. and Harlech-Jones, B. (2008). The fault is in ourselves: looking at ‘failures in implementation’. *Compare*, 38(1):39–51.
- Motola, S. (2001). Quality and indicators of quality in south african education: a critical appraisal. *International Journal of Educational Development*.
- Muralidharan, K. and Sundararaman, V. (2013). Contract teachers: experimental evidence from india. Technical report, National Bureau of Economic Research.
- Murnane, R. J. and Ganimian, A. J. (2014). Improving educational outcomes in developing countries : lessons from rigorous evaluations. *National Bueau of Economic Research*.

- Murnane, R. J., Willett, J. B., Duhaldeborde, Y., and Tyler, J. H. (2000). How important are the cognitive skills of teenagers in predicting subsequent earnings? *Journal of Policy Analysis and Management*, 19(4):547–568.
- NCERT (2011). What do they know? a summary of india’s national achievement survey, class v, cycle 3 2010/11. Technical report, National Council of Educational Research and Training.
- Nelson, R. R. and Phelps, E. S. (1966). Investment in humans, technological diffusion, and economic growth. *The American Economic Review*, pages 69–75.
- NEMIS-AEPAM (2011). An analysis of educational indicators of pakistan-2011. *AEPAM Publication*.
- Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, pages 1417–1426.
- Nye, B., Konstantopoulos, S., and Hedges, L. V. (2004). How large are teacher effects? *Educational evaluation and policy analysis*, 26(3):237–257.
- Osei, G. M. (2006). Teachers in ghana: issues of training, remuneration and effectiveness. *International Journal of Educational Development*, 26(1):38–51.
- Pitt, M. M., Rosenzweig, M. R., and Gibbons, D. M. (1993). The determinants and consequences of the placement of government programs in indonesia. *The World Bank Economic Review*, 7(3):319–348.
- Pritchett, L. (2013). *The Rebirth of Education: Schooling Ain’t Learning*. CGD Books.
- Rasch, G. (1961). On general laws and the meaning of measurement in psychology. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*, volume 4, pages 321–333. University of California Press Berkeley, CA.
- Rivkin, S. G., Hanushek, E. A., and Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2):417–458.
- Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, pages 247–252.
- Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics*, 125(1):175–214.
- Saeed, M. and Mahmood, K. (2002). Assessing competency of pakistani primary school teachers in mathematics, science and pedagogy. *International Journal of Educational Management*, 16(4):190–195.
- Schultz, T. W. (1961). Investment in human capital. *The American Economic Review*, 51(1):1–17.
- Todd, P. E. and Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485):F3–F33.
- Tymms, P. (1999). Baseline assessment, value-added and the prediction of reading. *Journal of Research in Reading*, 22(1):27–36.

- UNESCO (2013). Status of teachers in pakistan. Technical report, UNESCO and ITA.
- UNESCO (2014). Enseigner et apprendre : Atteindre la qualité pour tous. Technical report, United Nations Educational, Scientific and Cultural Organization.
- Verbeek, M. and Nijman, T. (1992). Testing for selectivity bias in panel data models. *International Economic Review*, pages 681–703.
- Verspoor, A. (2005). *The challenge of learning: Improving the quality of basic education in sub-saharan Africa*. Association for the Development of Education in Africa (ADEA).
- Westbrook, J., Shah, N., Durrani, N., Tikly, C., Khan, W., and Dunne, M. (2009). Becoming a teacher: Transitions from training to the classroom in the nwfp, pakistan. *International Journal of Educational Development*, 29(4):437–444.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.
- WorldBank (2013). Bangladesh education sector review. Technical report, World Bank, Human development Sector South Asia Region.

8 Annexes

A Annex - questions within the LEAPS tests

The LEAPS project tests children in grade three, four and five in three subjects: English, Urdu and Mathematics. The items selected in these tests aim to cover the general range of content taught to pupils by the time they complete the fifth grade. The difficulty and the types of questions (MQC, short or long answer) vary across the items in order to capture a large variation in achievement. Technical annexes in Andrabi et al. (2007) present in more details the characteristics of the tests used in the LEAPS survey. Andrabi et al. (2007) also provide evidence of the consistency of these tests.

English and Urdu

Contrary to other assessments, the items used in the LEAPS project do not make the assumption that pupils have the basic ability to read and write because the investigators in the pilot project observed that students were considerably below curriculum targets. For this reason, both English and Urdu tests begin with questions about alphabet, progress through more complicated elements of writing (word construction and recognition, grammar, vocabulary, sentence construction) and conclude with reading comprehension and an essay. The starting questions of the English test are easier than the Urdu test since knowledge in English tends to be lower.

Mathematics

The test in Mathematics identifies five major domains :

- Number recognition, properties and operations (40% of the test)
- Measurement (20% of the test)
- Geometry and spatial sense (15% of the test)
- Data analysis, statistics and probabilities (15% of the test)
- Algebra and functions (15% of the test)

The range of difficulty varies within the test (i-e addition with one digit, two digits and three digits, with decimal...) along with the skills tested (counting, percentages, fractions).

B Annex - Econometric issues

In this appendix, we address three main empirical challenges in order to estimate properly the gain model (equation 7) : attrition , selection, and assumptions of the gain model.

B.1 Attrition

The first potential econometric issue comes from the temporal dimension of our database even though, as underlined before, attrition is relatively low for a panel in a developing country. Children who drop out of the study may be systematically different from those remaining in the study leading to

attrition bias. Table B1 presents the descriptive statistics separately for the stayers and the attriters. We observe significant differences between these children. As expected, children who attrit have worst academic results and come from less educated and poorer households. Boys and elder children also tend to disappear from the study. Children who attrit also have specific characteristics in terms of school attended and teacher. Children attending a private school charging higher fees are more likely to drop out. Attriters live on average further to their schools. Having a female or a same-gender teacher is associated with less attrition. Children who attrit have on average teachers who are less paid and many of them have temporary contract teachers.

Table B1: Attrition: descriptive statistics

	No Attrition		Attrition		Difference	
	Mean	Sd	Mean	Sd	Diff	T-statistic
<i>Child characteristics</i>						
Std English IRT scores (MLE)	0.04	1.0	-0.11	1.1	0.15***	(7.22)
Std Math IRT scores (MLE)	0.05	1.0	-0.13	1.1	0.18***	(8.79)
Std Urdu IRT scores (MLE)	0.06	1.0	-0.16	1.1	0.22***	(10.84)
Child age	9.58	1.4	9.80	1.6	-0.22***	(-7.26)
Girl	0.47	0.5	0.40	0.5	0.07***	(6.46)
Child overweight	0.01	0.1	0.01	0.1	-0.00	(-0.29)
Child underweight	0.12	0.3	0.11	0.3	0.01	(1.33)
<i>Household characteristics</i>						
Dad uneducated	0.34	0.5	0.38	0.5	-0.04**	(-3.05)
Dad less primary	0.06	0.2	0.10	0.3	-0.03***	(-4.94)
Dad primary to high sec	0.53	0.5	0.46	0.5	0.07***	(5.28)
Dad more high sec	0.06	0.2	0.06	0.2	0.00	(0.37)
Mum uneducated	0.65	0.5	0.66	0.5	-0.00	(-0.41)
Mum less primary	0.07	0.2	0.08	0.3	-0.02*	(-2.21)
Mum primary to high sec	0.27	0.4	0.24	0.4	0.03*	(2.17)
Mum more high sec	0.01	0.1	0.02	0.1	-0.01	(-1.74)
Elder siblings (living in the hh)	2.42	2.1	2.22	2.1	0.20**	(3.23)
Wealth index	-0.30	1.4	-0.42	1.3	0.12**	(2.94)
<i>School characteristics</i>						
Private school	0.29	0.5	0.32	0.5	-0.03**	(-3.23)
15 min walking between home and school	0.75	0.4	0.69	0.5	0.06***	(4.54)
School has a library	1.67	0.5	1.68	0.5	-0.01	(-1.13)
School has computer facilities	1.90	0.3	1.89	0.3	0.01*	(2.28)
School has sport facilities	1.77	0.4	1.73	0.4	0.04***	(4.25)
School has an activity room	1.88	0.3	1.85	0.4	0.04***	(4.97)
School has four walls	1.22	0.4	1.24	0.4	-0.02*	(-2.28)
School has fans/rooms coolers	1.22	0.4	1.24	0.4	-0.02*	(-2.10)
School has electricity	1.17	0.4	1.19	0.4	-0.02**	(-3.03)
Admission fees (Grades 1-3)	37.77	115.2	43.88	132.8	-6.12*	(-2.45)
Annual school fees (Grades 1-3)	389.39	783.6	426.37	817.1	-36.98*	(-2.25)
Admission fees (Grades 4-5)	45.81	165.2	53.42	181.4	-7.60*	(-2.16)
Annual school fees(Grades 4-5)	475.11	936.8	529.07	1029.5	-54.00**	(-2.70)

Continued on next page

Following the previous table

<i>Teacher characteristics</i>						
Female Teacher	0.57	0.5	0.53	0.5	0.04***	(4.11)
Same gender Teacher	0.83	0.4	0.79	0.4	0.05***	(5.63)
Teacher experience	9.75	8.4	9.53	8.4	0.22	(1.22)
Teacher training	0.78	0.4	0.77	0.4	0.01	(1.62)
Teacher's years of education	11.48	2.1	11.52	2.2	-0.04	(-0.83)
Diploma : Matric	0.43	0.5	0.42	0.5	0.01	(1.27)
Diploma : MA or higher	0.06	0.2	0.06	0.2	0.00	(0.19)
Diploma : FA-FSc	0.30	0.5	0.31	0.5	-0.01	(-0.71)
Diploma : BA-BSc	0.21	0.4	0.21	0.4	-0.00	(-0.30)
Permanent contract	0.63	0.5	0.60	0.5	0.03**	(3.17)
Temporary contract	0.37	0.5	0.40	0.5	-0.03**	(-3.17)
Teaching wage (Rupees)	4197.79	2274.9	4082.22	2270.0	115.6*	(2.45)
Can receive bonus or prize	0.34	0.5	0.38	0.5	-0.05***	(-4.51)
Did receive bonus or prize	0.09	0.3	0.09	0.3	-0.01	(-0.95)
Teacher other work : agriculture	0.17	0.4	0.18	0.4	-0.01	(-1.34)
Teacher other work : business	0.03	0.2	0.04	0.2	-0.01	(-1.56)
Teacher other work : teaches out-side	0.10	0.3	0.13	0.3	-0.02***	(-3.36)
Teacher other work	0.04	0.2	0.04	0.2	0.01*	(2.12)
Teacher absence (days last month)	2.14	3.2	2.16	3.3	-0.02	(-0.28)
<i>Classroom characteristics</i>						
Class size	28.87	16.5	27.66	16.4	1.21***	(3.45)
% with English books	0.77	0.3	0.78	0.3	-0.00	(-1.09)
% with Math books	0.78	0.3	0.78	0.3	-0.01	(-1.07)
% with Urdu books	0.79	0.3	0.80	0.3	-0.01	(-1.49)
% with Desks	0.46	0.4	0.45	0.4	0.01	(0.84)
% with Chairs	0.16	0.4	0.17	0.4	-0.02*	(-2.06)
% with Blackboards	0.82	0.3	0.81	0.3	0.01	(1.96)
% girls in the class	0.46	0.4	0.41	0.4	0.06***	(6.57)
Observations	8338		3215		11553	
<i>Notes</i> : t statistics in parentheses : * $p < .05$, ** $p < .01$, *** $p < .001$						
Attrition represents children who were tested in the first wave but not in the third wave.						
<i>Source</i> : Author using the three waves of the LEAPS database, only children present in the first wave.						

A probit regression explaining the reasons for attrition (Table B2) is estimated to determine whether attrition is random (Fitzgerald et al., 1998). The probit model however shows that very few of our observable variables explain attrition. The pseudo R-squared obtained is relatively low suggesting that the proportion of attrition which is non-random is low. However we also find that higher scores tend to reduce the probability of dropping out.

Table B2: Probit : attrition test 1

Dep Var : Attrition = 1	(1)	(2)	(3)	(4)
Continued on next page				

Following the previous table

Std English Theta (MLE)	-0.024 (0.031)	-0.103*** (0.029)		
Std Math Theta (MLE)	-0.034 (0.030)		-0.113*** (0.026)	
Std Urdu Theta (MLE)	-0.105*** (0.028)			-0.138*** (0.024)
Child age	0.055*** (0.012)	0.055*** (0.012)	0.056*** (0.012)	0.054*** (0.012)
Girl	-0.022 (0.058)	-0.032 (0.060)	-0.045 (0.060)	-0.017 (0.059)
Dad uneducated	0.095 (0.090)	0.105 (0.090)	0.119 (0.090)	0.097 (0.090)
Dad less primary	0.312*** (0.111)	0.329*** (0.110)	0.335*** (0.110)	0.315*** (0.110)
Dad primary to high sec	-0.036 (0.077)	-0.030 (0.076)	-0.018 (0.077)	-0.036 (0.077)
Mum uneducated	-0.408** (0.169)	-0.403** (0.168)	-0.405** (0.171)	-0.406** (0.169)
Mum less primary	-0.150 (0.182)	-0.144 (0.180)	-0.152 (0.183)	-0.145 (0.181)
Mum primary to high sec	-0.347* (0.180)	-0.345* (0.179)	-0.345* (0.181)	-0.346* (0.180)
Elder siblings (living in the hh)	-0.025*** (0.009)	-0.025*** (0.009)	-0.024*** (0.009)	-0.025*** (0.009)
Wealth index	-0.030** (0.015)	-0.034** (0.015)	-0.035** (0.015)	-0.030** (0.015)
Less 15 min to go to school	-0.383*** (0.076)	-0.388*** (0.076)	-0.377*** (0.075)	-0.384*** (0.076)
15-30 min to go to school	-0.323*** (0.077)	-0.326*** (0.077)	-0.322*** (0.076)	-0.323*** (0.076)
Female Teacher	-0.139** (0.057)	-0.128** (0.058)	-0.148** (0.058)	-0.134** (0.057)
Same gender Teacher	-0.124* (0.073)	-0.129* (0.073)	-0.127* (0.074)	-0.124* (0.073)
Teacher Experience	0.001 (0.005)	0.000 (0.005)	0.001 (0.005)	0.001 (0.005)
Teacher Education	0.015 (0.010)	0.018* (0.010)	0.012 (0.011)	0.014 (0.011)
No training	-0.100 (0.087)	-0.083 (0.085)	-0.104 (0.083)	-0.100 (0.086)
Teacher absence (days last month)	0.003 (0.009)	0.004 (0.009)	0.003 (0.009)	0.003 (0.008)
Permanent contract	-0.028 (0.066)	-0.022 (0.066)	-0.031 (0.066)	-0.029 (0.066)
Log of teacher wage	-0.001 (0.038)	0.004 (0.037)	0.001 (0.037)	-0.002 (0.037)
Class size	-0.005*** (0.002)	-0.005*** (0.002)	-0.005*** (0.002)	-0.005*** (0.002)
Private school	0.152 (0.173)	0.150 (0.171)	0.120 (0.170)	0.137 (0.174)
Log of annual fees (grades 1 to 3)	0.017	0.016	0.017	0.016

Continued on next page

Following the previous table

	(0.020)	(0.020)	(0.020)	(0.020)
Less 15 min from school to reach health center	0.136*	0.140*	0.130*	0.131*
	(0.077)	(0.075)	(0.078)	(0.077)
15-30 min from school to reach health center	0.018	0.022	0.020	0.018
	(0.078)	(0.076)	(0.077)	(0.077)
Constant	-0.422	-0.496	-0.414	-0.400
	(0.462)	(0.458)	(0.462)	(0.452)
Observations	5312	5312	5312	5312
Pseudo R ²	0.037	0.032	0.033	0.036

Notes : robust clustered (by village) errors in parentheses : * $p < .05$, ** $p < .01$, *** $p < .001$

Attrition represents children who were tested in the first wave but not in the third wave.

Source : Author using the first wave of the LEAPS database.

In the specific case of fixed effects models, attrition and sample selection are only a problem if selection is related with the idiosyncratic errors (Wooldridge, 2010). To test this, the method suggested by Verbeek and Nijman (1992) is implemented. Verbeek and Nijman (1992) suggest a test in the context of random effects but Wooldridge (2010) proves that it can be used in fixed effects models. We add a variable, the lead selection indicator ($s_{i,t+1}$), to the contemporaneous model (equation 6).¹⁵ This variable equals zero if the child does not attrit and changes from zero to one in the period just before the student attrits. For this test, we exclude children who were surveyed in round one, disappeared in round two and reappeared in the last round. If idiosyncratic errors are uncorrelated with attrition, the lead selection variable indicating attrition in $t + 1$ should not impact achievement in t . The results presented in table B3 suggest that attrition is not related with idiosyncratic errors.

Table B3: Attrition test 2 : Verbeek and Nijman (1992) test

Dep. Var.	(1) Std English score (IRT)	(2) Std Math score (IRT)	(3) Std Urdu score (IRT)
Lead selection indicator	0.028 (0.040)	-0.033 (0.041)	0.004 (0.042)
Observations	17097	17097	17097
Adjusted R ²	0.028	0.023	0.021
FE Schools	Yes	Yes	Yes
FE Teachers	No	No	No
FE Students	Yes	Yes	Yes

Notes : Robust clustered standard errors in parentheses : * $p < .05$, ** $p < .01$, *** $p < .001$

Control variables not presented : same variables presented in table 6.

Source : Author using the three waves of the LEAPS database.

Given that attrition is relatively low in our sample and tests provide evidence that estimates with student and school fixed effects are not biased, we use the unbalanced panel sample for our main estimates. We nevertheless provide the results for the gain model using the balanced sample as a

robustness check.

As underlined by Harris and Sass (2011); Kane et al. (2008) and Rockoff (2004), the estimated effects of observed teacher characteristics could also be biased if teacher attrition is not accounted for. If more effective teachers are less likely to leave the profession, it could lead to upward bias in estimates of teacher effects. The bias could also be downward if more effective teachers have higher opportunity costs or better job opportunities outside of teaching and are therefore more likely to leave teaching profession. Around 15% of the teaching staff left their schools during the two previous years. However, most of the teachers (59%) left their previous schools because of personal and family issues (marriage, domestic problems, illness or going to have further education) and only 17% joined another school. In order to assess if teachers who attrit have different characteristics than those who stay in the profession, Table B4 presents the results of t tests on the equality of means between attriting teachers and other teachers. Local women teachers and less experienced, trained and educated teachers are over-represented in the attriters' group. Table B5 presents the results from a probit explaining teacher attrition by teacher observable characteristics. Low levels of experience, education, absenteeism and temporary contracts explain partly teacher attrition.

Table B4: Mean-comparison tests between teachers who stay in the school and teachers who attrit

	Teachers who do not attrit - teachers who attrit	
	Mean diff.	T-statistics
Female Teacher	-0.14***	-7.54
Age of teacher	6.36***	16.37
Teaching experience < 1 year	-0.16***	-12.95
Teaching experience 1-3 years	-0.16***	-10.52
Teaching experience > 3 years	0.32***	18.60
Teaching experience in current school < 1 year	-0.22***	-14.51
Teaching experience in current school 1-3 years	-0.09***	-5.66
Teaching experience in current school > 3 years	0.31***	17.18
Teacher diploma : Matric or less	0.01	0.59
Teacher diploma : FA/FSc	-0.06***	-3.58
Teacher diploma : BA/BSc	-0.01	-0.47
Teacher diploma : Master or more	0.06***	4.59
Teacher training : No	-0.29***	-16.15
Teacher training : PTC	0.17***	9.65
Teacher training : CT	0.07***	4.66
Teacher training : B.Ed	0.05***	3.46
Teacher contract : permanent	0.42***	23.34
Teacher contract : temporary	-0.42***	-23.34
Local teacher	-0.06***	-3.37
Monthly salary	1973.53***	17.18
Days teacher did not attend class last month?	-0.18	-1.01
Observations	4760	

Continued on next page

Following the previous table

Notes : * $p < .05$, ** $p < .01$, *** $p < .001$

Teachers who attrit are teachers who left their schools whatever the reason for leaving. Statistics represent characteristics of the teachers in the first year, when they all were present in the sample. Teachers are all teachers in the school (all grades) not only teachers of the tested students.

Source : Author using the three waves of the LEAPS database.

Table B5: Probit : attrition of teachers

Dep. var. : attrition=1	(1)
Female Teacher	0.053 (0.054)
Age of teacher	0.007 (0.004)
Teaching experience < 1 year	0.057 (0.108)
Teaching experience 1-3 years	0.046 (0.101)
Teaching experience in current school < 1 year	0.348*** (0.096)
Teaching experience in current school 1-3 years	0.148* (0.082)
Teacher diploma : matric or less	0.093 (0.120)
Teacher diploma : FA/FSc	0.222** (0.103)
Teacher diploma : BA/BSc	0.161* (0.093)
Teacher training : no	0.055 (0.114)
Teacher training : PTC	-0.189* (0.104)
Teacher training : CT	-0.186* (0.106)
Teacher contract : permanent	-0.901*** (0.096)
Local teacher	-0.022 (0.047)
Monthly salary	0.000 (0.000)
Days of absence last month	0.025*** (0.005)
Constant	-1.116*** (0.155)
Observations	4521
Pseudo R ²	0.1377

Continued on next page

Following the previous table

Notes : Robust clustered (by village) errors in parentheses : * $p < .05$, ** $p < .01$, *** $p < .001$

Attrition represents teachers who leave their schools for any reason in rounds 2 or 3.

Reference categories : Male teacher with a Master diploma or more and a temporary contract. He has three years of total experience and experience in his current school or more. He followed a B.Ed training program.

Source : Author using the first wave of the LEAPS database.

To control for teacher attrition, following Hanushek et al. (2005); Harris and Sass (2011) and Rockoff (2004), we also provide estimates including both student and teacher fixed effects. As pointed out by Harris and Sass (2011), the inclusion of teacher fixed effects reduces the potential bias associated with teacher attrition but it does not completely eliminate it. If unobserved time-varying teacher characteristics are correlated with the probability of attrition, it would not be fully captured by teacher fixed effects. We will also provide the estimates using the balanced sample with only the teachers surveyed during all the waves.

B.2 Selection and endogeneity

Besides students and teachers attrition, if students, school resources and teachers are not assigned to schools and classrooms randomly, the specifications presented before could be biased by endogeneity (Ishii and Rivkin, 2009). Non-random assignment consists of four main components : sorting of students to schools, sorting of students to teachers within schools, sorting of teachers to schools and sorting of teachers within contracts. In this subsection, we explore the extent to which these four types of selection could bias our estimates.

First, if parents choose the school attended by their children because of its alleged quality (Pitt et al., 1993) or if the school selects children it might bias the results. School choice faced by the parents appears to be relatively important (Table B6). A typical village in our sample has eight schools, five of which are within 50-100 meters of each other (Andrabi et al., 2007). Despite this relative large schooling offer, parental choice is mainly driven by distance and budgetary constraints whereas the quality of the school is not a major determinant (Table B7).

Table B6: Competition within schools

	All schools	Public schools	Private schools
Government schools <15 min walk	1.66	1.21	2.38
Private schools <15 min walk	1.87	1.52	2.45
Madrassas < 15min walk	0.94	0.70	1.32
All schools < 15min walk	4.47	3.43	6.14

Lecture : On average, public schools have 3.43 other schools and 1.21 public schools located within 15 minutes walk.

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

Non-random assignment of students could also come from the schools themselves. However, in our sample, the selection of students by schools appears to be low. If 81% of the children attend a school where there is a specific procedure for admitting pupils, most of the schools (98%) admit every student who applied. So in theory, schools do apply criteria to select students but, in practice, they admit almost everyone. The inclusion of school fixed effects partially deals with non-random time-invariant assignment of students within schools. However, if parents respond dynamically to

Table B7: School choice : reasons for attending this school

	First reason %	Second reason %
Close to Home	42.59	8.18
Low Cost	26.68	44.95
High Quality	20.79	25.08
No other option	4.03	11.58
Relative/friend owns/teaches in school	2.23	4.58
Female teachers	0.20	0.23
Male teachers	0.00	0.23
Children from same Biradari go here	1.22	1.96
Observation (hh)	2954	2556

Source : Author using the first wave of the LEAPS database.

changes in school quality by transferring their children to other schools, school fixed effects would not correct for this bias. Evidence from the data suggests that is not a relevant issue. Indeed, even when the child switched to another school, which is very rare (less than 2% have changed between each round), the main reasons were because the school was too expensive or simply because the class was not offered and not because of low quality (Table B8).

Table B8: School choice : reasons for leaving school

	First reason %	Second reason %
Classes not offered	23.44	36.31
Too expensive	17.18	13.77
Low quality	18.54	11.09
Too far	6.60	2.68
Family moved	3.30	2.50
For religious education	1.82	4.47
Relative/friend owns/teaches in other school	1.59	1.43
Punishment from teachers	1.48	3.22
No male teachers	0.80	0.54
No female teachers	0.46	0.18
Other	24.80	23.79
Observation (hh)	879	559

Source : Author using the first wave of the LEAPS database.

Besides, parents appear to have very few information about the quality of the diverse available schools (Table B9). However, because few students have left schools due to school quality reasons, as a robustness check, we estimate the gain model with student and school-by-year fixed effects (instead of simple school dummies).

Our estimates could also be biased if, within a school, students are assigned to specific teachers whereas it is on the demand of their parents or of the head teacher. We assume that parents have no way to pressure school principals in order to place their children in a class with a better teacher or to get the school to invest more in their children's education. This assumption seems to be relevant with our data as parents do not have a good idea about teacher quality. Indeed, only 53% of adults' members in the households know directly the name of their child's teachers, 57% do not know the educational qualifications of the teachers and 13% cannot tell if he was absent last month or not or if the teacher is good or not. Parents could in theory have an influence on school inputs if they were

Table B9: Parental information about schooling offer

	Female member	Male member
Have you heard about this school?	0.56	0.69
Can you visit this school whenever you want?	0.71	0.74
Have you ever been to this school?	0.28	0.33
Don't know quality of school (%)	0.41	0.47
Quality of school : poor or very poor (%)	0.08	0.06
Quality of school : average (%)	0.31	0.25
Quality of school : above average (%)	0.19	0.22
Observation (hh)	13335	13358

Source : Author using the first wave of LEAPS database.

involved in school decision making processes through school committees or parent-teacher associations. However, only 7% and 13% of respectively men and women have ever participated in school committees or parent-teacher meetings. Moreover, even when they participate in school committees, in 68% of our sample schools, school-committees or parent-teacher associations do not play any role in selecting and promoting teachers. For these reasons, school inputs are assumed to depend only upon school's staff and not on parents. If the parents do not choose the school or the teacher, school principals could however assign students to specific teachers leading to potential endogeneity issues. Evidence from developed countries tends to show that better trained and experienced teachers are assigned to students with greater ability (Clotfelter et al., 2005; Feng, 2009). If children with higher ability are systematically assigned to the best students, teacher effects will be biased upward. To test if there are systematic differences of students assigned to particular teachers, we test if current classrooms are significant predictors of past test scores. Following Rockoff (2004), we proceed in two steps. First, the residuals from a regression of past scores on school-year dummies are estimated. Then, we regress these residuals on teachers' dummies and use a joint F-test to test the significance of the coefficient. Table B10 shows, by subject, the F-statistics and p-values of these tests. As all the p-values are close to one, we can assume that, within a school, there is no specific classroom assignment based on achievement. In order to control for student-teacher selection, studies in the USA have made use of experiments with random assignment of students and teachers or natural experiments (Clotfelter et al., 2006; Dee, 2004; Garet et al., 2008, 2010; Jacob and Lefgren, 2004; Nye et al., 2004). In the absence of such experiments, in this paper, following studies in USA (Betts et al., 2003; Boyd et al., 2006; Clotfelter et al., 2007; Hanushek et al., 2005; Jepsen, 2005; Rivkin et al., 2005; Rockoff, 2004), we use student fixed effects to control for student heterogeneity and student-teacher matching. Student fixed effects alleviate the bias due to students assignment to teachers based on time-invariant student characteristics (static tracking) (Koedel and Betts, 2011; Rothstein, 2010).

Table B10: Test for systematic classroom assignment

	F-statistic	P-value
English	0.24	1.00
Mathematics	0.27	1.00

Note : F-tests represent the joint significance of teachers' dummies to predict past scores within school-year cells.

However, if teacher assignments are correlated with time-varying error terms, it invalidates the value added model (equation 11) even if there is no static tracking and if student fixed effects are included (Koedel and Betts, 2011; Rothstein, 2010). Students can be sorted dynamically based on time-variations (or shocks) to test-score-growth trajectories. For instance, students for an unobserved reason, get an unusually high score in year $t - 1$ and then they fall back to their typical gain score. In time t , if students are assigned dynamically based on their shocks in scores, their assignment to teacher in time t will be correlated to this shock which is captured by $e_{i,t-1}$. Teacher effects in equation 11 cannot be interpreted as causal relationships if $e_{i,t}$ are serially correlated. Indeed T_{it} will be correlated with shocks in $t - 1$ which are correlated with shocks in t captured by $e_{i,t}$. Therefore, as underlined by Harris and Sass (2011) and Rothstein (2010), the three-way fixed effects approach, presented in equation 12, reduces bias associated to students and teachers sorting but could still be biased if students are dynamically assigned to teachers based on prior unobserved shocks which are serially correlated. We conduct an exogeneity falsification test as suggested by Rothstein (2010) and Koedel and Betts (2011). This procedure, by testing for effects of future teachers on current achievement gains, determines whether value-added models can provide causal information about teaching effectiveness. Rothstein (2010) adds future teacher assignments into the regression on current test scores. Because there is no reason why future teachers could have causal effects on current achievement, if we observe an effect of teacher assignments on current scores it probably results from a correlation between teacher assignments and the error terms. On the contrary, if the coefficients on future-teacher dummies are jointly insignificant, it suggests that the controls in the model capture the sorting bias that would have been confounded with teacher effects. As proposed by Koedel and Betts (2011), we test for effects of future teachers on current achievement gains using the following model :

$$A_{i4} - A_{i3} = \gamma^3 T_{i3} + \gamma^4 T_{i4} + \gamma^5 T_{i5} + s_m + \epsilon_i \quad (14)$$

where $A_{i4} - A_{i3}$ refers to the gain score between grade three and four, s_m represents school fixed effects, T_{ix} is the vector of teacher indicator variables for student i in grade x and γ^x refers to the vector of teacher effects corresponding to teachers who teach in grade s . According to Rothstein (2010), if future teacher effects (fifth grade teachers in our case), measured by γ^5 , are jointly significantly different, teacher effects cannot be interpreted as causal. In order to replicate Rothstein (2010) test, we exclude students who change schools across grades and we focus on one single cohort, using fourth grade students. Our results (Table B11) confirm Rothstein's suspicions that future teachers explain a portion of current grade achievement gains. Our baseline model with school and student fixed effects could therefore be biased by dynamic assignment.

Table B11: Test for dynamic sorting (Rothstein, 2010; Koedel and Betts, 2011)

Dep. Var	Gain in English std IRT score		Gain in Math std IRT score		Gain in Urdu std IRT score	
	Wald Statistic	P-Value	Wald Statistic	P-Value	Wald Statistic	P-Value
Grade 4 Teachers	F(32,2439)=198.54	<0.01	F(32,2439)=39.30	<0.01	F(32,2439)=1005	<0.01
Grade 5 Teachers	F(51,2439)=12.94	<0.01	F(53,2439)=38.63	<0.01	F(50,2439)=11972	<0.01

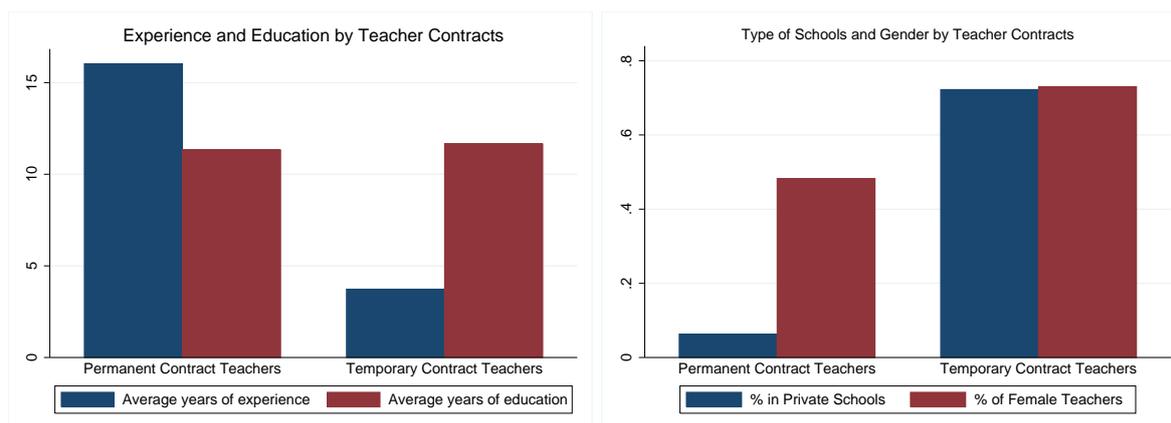
Source : Author using student gain in grade 4 and past, current and future teacher dummies along with current school fixed effects. The Wald Statistics and p-values refer to tests that all teachers in the given grade have identical effects on students' gains in grade 4. Only children who changed teachers between grade 4 and 5.

Literature suggests that more qualified and experienced teachers tend to choose schools with higher achieving students coming from wealthier families (See Boyd et al. (2005); Clotfelter et al. (2005); Hanushek et al. (2004); Horng (2009) for studies in the United States). If teachers choose their schools based on unobserved characteristics, estimates can be biased. The same thing occurs if teachers choose, within a school, the classroom and the students they have. Teachers can theoretically choose their schools (Hanushek and Rivkin, 2006) but we assume it is not the case in our model. Indeed in Pakistan, the Provincial Directorate of Education appoints all government primary school teachers. Moreover, in Punjab, the local government policy stipulates that 75% of the recruited teachers must come from the area where the school is located. Andrabi et al. (2011) using the same database show that better teachers are not allocated to richer schools or villages on the basis of observed educational and professional qualifications. School fixed effects are nevertheless included to reduce the bias due to selection of teachers within schools.

Biases could also emerge due to non-random selection of teachers into contracts. If teacher contracts are correlated with other unmeasured teacher characteristics, estimates of the impact of contract would be biased. In the case of teacher fixed effects, biases would exist if teacher contracts are correlated with time-varying unobservable characteristics. For instance, if teachers exhibiting low productivity are assigned to temporary contracts, the estimated effects of teacher contracts would be biased downward. The assignment of teachers to contract is only problematic if unmeasured factors are correlated with both contract selection and teacher performance. Figure B1 provides histograms of teacher contract by level of schooling, experience, gender and type of schools. They show a clear pattern of teachers being predominantly hired with temporary contracts in private schools. As school fixed effects are included in all our regressions, this correlation cannot be the reason of biases due to teacher selection into contracts. Teachers are also less hired with temporary contracts as they gain experience, a variable also included in all the regressions.

To better understand the factors driving selection into contracts, we conduct a probit analysis where the dependent variable equals one if the teacher has a temporary contract. To account for possibility that having a temporary contract is driven by past teacher performance, which is the main potential source of bias due to contract selection, we include average scores of the teacher's students in the previous year. Other explanatory variables include teacher education, experience, gender, training, geographical background and finally district and year fixed effects. The results presented in table B12 show that selection within contracts does not depend on teacher past performances

Figure B1: Teacher Contracts by experience, education, gender and types of schools



Source : Author using the three waves of the LEAPS project

reducing the risk of bias. Selection within contracts seems to depend highly on experience and types of school which we control for.

Table B12: Probit estimates of teacher contracts

Dep var : =1 if temporary contract teachers	(1)	(2)	(3)	(4)	(5)	(6)
Lagged average English scores of teacher's students	0.012 (0.098)			-0.015 (0.100)		
Lagged average Math scores of teacher's students		-0.127 (0.098)			-0.110 (0.102)	
Lagged average Urdu scores of teacher's students			-0.062 (0.094)			-0.056 (0.100)
Teacher Experience	-0.133*** (0.023)	-0.132*** (0.023)	-0.132*** (0.023)	-0.114*** (0.023)	-0.113*** (0.023)	-0.113*** (0.023)
Teacher Education	0.091* (0.055)	0.096* (0.055)	0.095* (0.055)	0.035 (0.059)	0.039 (0.059)	0.039 (0.060)
Female Teacher	0.064 (0.187)	0.054 (0.183)	0.084 (0.186)	0.094 (0.187)	0.079 (0.187)	0.105 (0.187)
Local Teacher	-0.131 (0.180)	-0.129 (0.179)	-0.129 (0.179)	-0.114 (0.189)	-0.112 (0.188)	-0.113 (0.188)
Private School	1.498*** (0.249)	1.576*** (0.252)	1.548*** (0.247)	1.375*** (0.292)	1.420*** (0.294)	1.394*** (0.288)
No training				-0.041 (0.338)	-0.037 (0.336)	-0.031 (0.341)
PTC training				-0.857*** (0.220)	-0.847*** (0.223)	-0.850*** (0.223)
CT training				-0.298 (0.238)	-0.293 (0.241)	-0.297 (0.240)
Constant	-0.632 (0.770)	-0.811 (0.764)	-0.767 (0.776)	0.257 (0.812)	0.118 (0.803)	0.158 (0.820)
FE districts	Yes	Yes	Yes	Yes	Yes	Yes

Continued on next page

Following the previous table

FE years	Yes	Yes	Yes	Yes	Yes	Yes
Observations	822	822	822	821	821	821
Pseudo R ²	0.562	0.563	0.562	0.585	0.586	0.585

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

Reference categories : Male teacher who has followed a B.Ed training program

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

To sum up, children sorting to schools is not likely to bias our estimates, neither is teacher selection of schools. To be sure to alleviate these biases, school fixed estimates are included in all regressions. Teacher selection into contract is not a source of major bias as it depends on variables we control for. Systematic matching to teachers is not a relevant concern and the inclusion of student fixed effects alleviates this bias. The only source of potential bias comes from student dynamical matching to teachers.

B.3 Assumptions for the value-added model

In section ??, we presented the different assumptions needed for the gain model to be valid. To test for the effects of the different inputs to be constant overtime, we include interactions between years and inputs in our gain model with student and school fixed effects. Table B13 presents Wald tests on each set of input interactions. Most of the inputs provide the same effects across years. Some exceptions have however to be made. For instance, teacher experience affects more students' achievement of younger cohorts.

Table B13: Test Value-added model 1 : Wald test of equality of interaction variables (inputs and years)

Dep. Var :	English IRT std score		Math IRT std score		Urdu IRT std score	
	F statistics	P value	F statistics	P value	F statistics	P value
Underweight	0,51	0,60	0,19	0,83	0,13	0,88
Overweight	2,05	0,13	0,21	0,81	0,13	0,87
Wealth index	0,25	0,78	0,21	0,81	1,02	0,36
Female Teacher	1,30	0,27	0,07	0,93	0,32	0,73
Same gender Teacher	0,30	0,74	1,38	0,25	1,11	0,33
Local Teacher	0,78	0,46	0,04	0,96	1,08	0,34
Teacher Exp	4,05	0,02	7,08	0,00	0,65	0,52
Teacher Educ	0,81	0,44	0,36	0,70	0,27	0,77
Non permanent contract teacher	2,43	0,09	0,67	0,51	0,65	0,52
Teacher training : PTC	4,45	0,01	4,17	0,02	0,52	0,59
Teacher training : CT	4,32	0,01	1,71	0,18	2,89	0,06
Teacher no training	1,77	0,17	1,78	0,17	4,02	0,02
Teacher log monthly wage	4,48	0,01	2,42	0,09	0,36	0,7
Bonus for students' perf	3,11	0,04	17,25	0,00	17,34	0,00
Other Bonus	0,78	0,46	1,44	0,24	0,01	0,99
Teacher Absenteism	0,92	0,40	1,15	0,32	1,28	0,28
Teacher teaches outside	2,35	0,10	1,06	0,35	2,87	0,06
Teacher other work	3,86	0,02	1,89	0,15	0,83	0,44
Class size	16,01	0,00	11,86	0,00	15,61	0,00
% with book in Eng/Math or Urdu	3,25	0,04	2,71	0,07	7,15	0,00

Continued on next page

Following the previous table

% with desk	3,16	0,04	4,29	0,01	4,94	0,01
% with chair	5,50	0,00	2,36	0,09	3,82	0,02
% Blackboards	4,78	0,01	0,66	0,52	0,21	0,81
% girls	2,01	0,13	1,04	0,35	1,08	0,34
Observations in the regression	13973		13973		13973	
Adjusted R ²	0.149		0.190		0.187	
FE Schools	Yes		Yes		Yes	
FE Teachers	No		No		No	
FE Students	Yes		Yes		Yes	

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

Source : Author, using the three waves of the LEAPS database

We also test for the influence of child's past achievement over current inputs by regressing one by one each input on past child's scores and other control variables (Table B14). The results tend to validate the assumption according to which parents and schools do not respond to students' past achievement.

Table B14: Test Value-added model 2

Dependent Variables	Coeff on lagged IRT std score					
	English	se	Math	se	Urdu	se
Female Teacher	0.014	(0.056)	-0.096*	(0.058)	-0.004	(0.051)
Same Gender Teacher	-0.003	(0.066)	0.198***	(0.067)	-0.108*	(0.065)
Local Teacher	-0.034	(0.055)	-0.060	(0.052)	0.013	(0.053)
Teacher Experience	-0.305	(0.349)	0.376	(0.322)	0.169	(0.306)
Teacher Education	0.089*	(0.051)	-0.016	(0.049)	-0.053	(0.047)
Teacher Temporary Contract	-0.004	(0.060)	-0.156***	(0.058)	-0.016	(0.058)
Teacher No Training	0.070	(0.086)	-0.075	(0.076)	0.017	(0.062)
Teacher (Log) Monthly Wage	0.008	(0.019)	0.054***	(0.020)	0.029*	(0.016)
Teacher Eligible for Bonus	-0.013	(0.053)	-0.044	(0.054)	-0.058	(0.054)
Teacher Absence	-0.086	(0.136)	-0.196	(0.154)	0.053	(0.096)
Class Size	0.120**	(0.060)	0.136**	(0.054)	0.105*	(0.057)
Teacher Other Work	-0.814***	(0.302)	-0.382	(0.264)	-0.455	(0.305)
% with English books	0.000	(0.006)	-0.001	(0.006)	-0.003	(0.005)
% with Math books	0.000	(0.006)	-0.002	(0.006)	-0.003	(0.005)
% with Urdu books	-0.000	(0.006)	-0.002	(0.006)	-0.003	(0.005)
% with Desks	0.015	(0.018)	0.004	(0.018)	-0.005	(0.017)
% with Chairs	0.001	(0.015)	-0.012	(0.014)	-0.007	(0.014)
% with Blackboards	-0.010	(0.012)	-0.005	(0.012)	-0.007	(0.011)

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

Control variables : child gender, dummies of household wealth, education of the mother and the father, private school dummy, number of teachers in the school, number of students in the school, dummies indicating whether the school has a library, computer facilities, sports facilities, four walls and electricity, time taken to go from school to the nearest telephone facility, bank, health center, public transport and districts fixed effects.

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

C Annex - Item Response Theory

Item Response Theory (IRT) is widely used in education fields since the 1970s' and 1980s' (Baker and Kim, 2004). It is for instance the method used to calculate PISA and TIMMS scores. Contrary to the Classical Test Theory (CTT), it does not assume that each item of a test is equally difficult. Two students who answer the same number of items will not be scored identically unless they have answered the same set of items correctly. In Item Response Theory, every item is distinct and given an ability level, the probability to get the right answer is different from one question to another because for instance one question is more difficult. The underlying of the Item Response Theory is that the probability of answering correctly to an item is a mathematical function of both individual and item parameters. The likelihood of being right to a specific question depends on individual latent ability (θ) and also on three different item parameters : difficulty (δ), discrimination (a) and pseudo-guessing (c) parameters. For each level of ability, IRT associates a certain probability of answering correctly the item using logistic distributions. The estimated scores take into account not only the number of questions answered correctly but also the types of question answered (difficulty, discrimination and guessing). Item Characteristics Curves (ICC) are graphical representations of the probability of being right in function of ability.

Three different logistic IRT models are generally used in educational studies. The first model, known as the one-parameter IRT model or the Rasch model (Rasch, 1961), implies that the probability for the individual i of being correct on item j , $P(\theta_i)$, is a logistic function of the difference between his latent ability (θ_i) and the item difficulty parameter (δ_j). If child's ability is greater than the difficulty of the question, it increases his probability of getting the correct response.

$$P(\theta_i) = \frac{1}{1 + e^{-1(\theta_i - \delta_j)}} \quad (15)$$

Figure C1 presents the Item Characteristic Curves generated for three different items with respectively difficulty parameters of 1, 0.5 and 0.2. The more difficult is the question, the lower the probability that a student with a certain given ability level will obtain a correct response. Increasing the difficulty of an item will shift the ICC to the right : an individual at the same level of ability will have a lower probability of answering right to the question.

In a two-parameter IRT model, the probability for the individual i of being correct on item j , $P(\theta_i)$, depends on the child's latent ability (θ_i), the item difficulty (δ_j) and also the item discrimination parameter (a_j).

$$P(\theta_i) = \frac{1}{1 + e^{-a_j(\theta_i - \delta_j)}} \quad (16)$$

The item discrimination parameter describes how well an item distinguish between individuals with the same ability level. It measures the slope of the characteristic curve at the point of inflection. Figure C2 presents the Item Characteristic Curves generated for three different items with respectively difficulty parameters of 1, 0.5 and 0.2 and discrimination parameters of 1, 1.2 and 0.9. The greater the slope of the ICC for the second item indicates that this item is more discriminating.

The two previous IRT models neglects the fact that individuals can answer correctly an item just

Figure C1: Item Characteristic Curves - One-parameter IRT Model

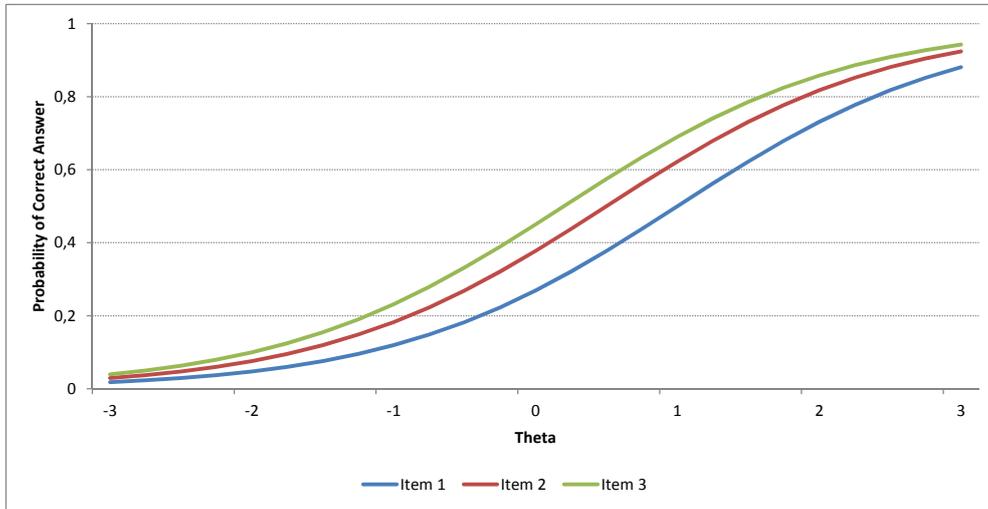
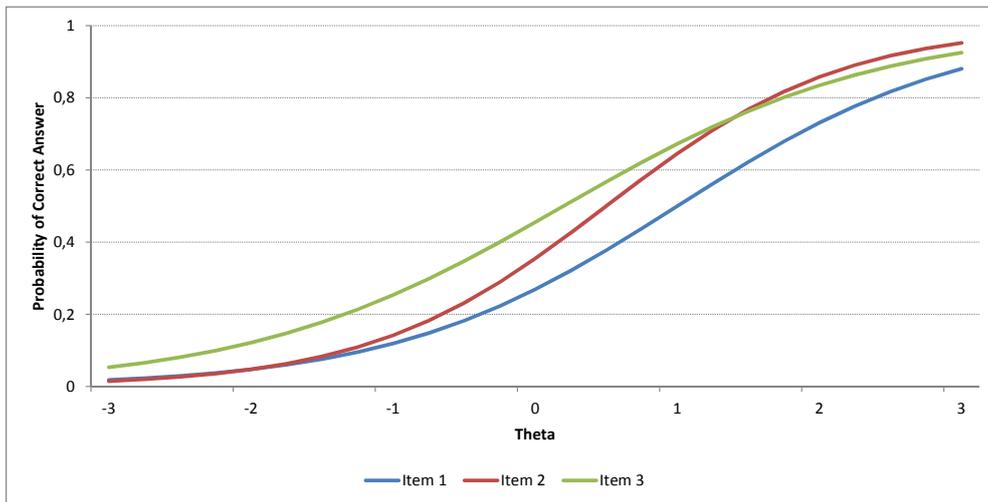


Figure C2: Item Characteristic Curves - Two-parameter IRT Model

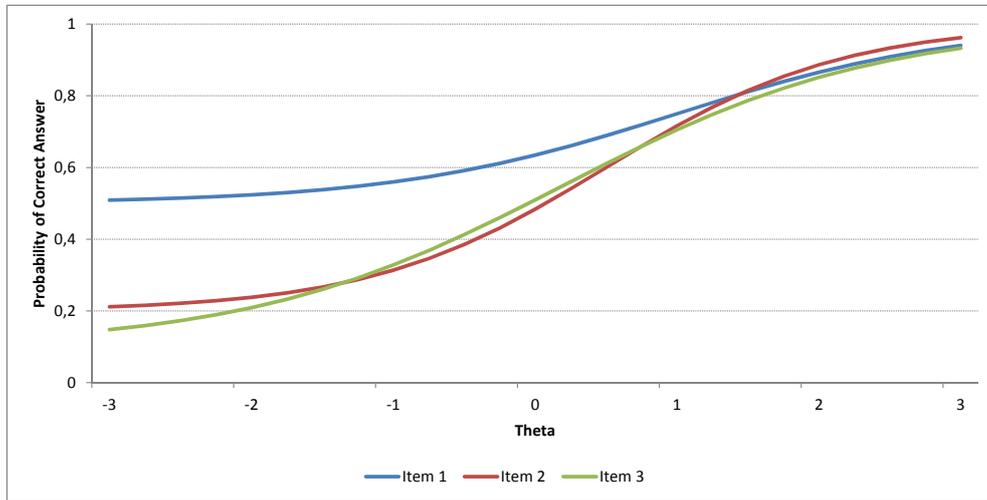


by guessing. The three-parameter IRT models the probability for the individual i of being correct on item j , $P(\theta_i)$, as a function of individual latent ability (θ_i), the item difficulty (δ_j), discrimination (a_j) and pseudo-guessing (c_j) parameters.

$$P(\theta_i) = c_j + \frac{1 - c}{1 + e^{-a_j(\theta_i - \delta_j)}} \quad (17)$$

The pseudo-guessing item parameter is simply the chances of being right just by guessing. This parameter ranges from zero to one and does not vary according to ability : individual with high or low ability have the same probability of guessing. Figure C3 shows the Item Characteristic Curves generated for three different items with respectively difficulty parameters of 1, 0.5 and 0.2, discrimination parameters of 1, 1.2 and 0.9 and pseudo-guessing parameter of 0.5, 0.2 and 0.1. The guessing parameter sets a floor to the lower value possible of the probability of answering right to the item.

Figure C3: Item Characteristic Curves - Three-parameter IRT Model



In this paper, the three-parameter IRT model is used to calculate scores. The three item parameters are estimated using maximum likelihood estimation. These parameters are computed so that they minimize the differences between the item characteristic curve and the observed proportions of correct responses from the data. Once these parameters are estimated, ability estimates are obtained by maximizing the likelihood function. This likelihood function describes the probability of each response pattern given the level of ability of the individual. A person taking a test with n items can have $n + 1$ scores (0,1,...,n) and the number of possible response to the test (response pattern) amounts to 2^k . The Characteristics Curves describe the probability of each response to each item but in order to calculate the probability of response patterns, we need to calculate the joint probabilities. Item Response Theory estimates of abilities are based upon one assumption : the local independence of items. Given an ability level, the answers to separate items are mutually independent. Because of this conditional independence assumption (conditional on θ_i), we can multiply probabilities of each items to get the probability of the whole pattern. The likelihood function for n items test is defined

as :

$$L(\theta) = \prod_{k=1}^n P_k(\theta_i, \delta_k, a_k, c_k)^{u_i} Q_k(\theta_i, \delta_k, a_k, c_k)^{1-u_i} \quad (18)$$

where u_k is the observed score at the k^{th} item and $Q_k = 1 - P_k$ is the probability, given the level of ability, of being wrong at the k^{th} item. Ability estimates, $\hat{\theta}_i$, are abilities with the highest likelihood given the observed pattern and the item parameters. The estimated scores take into account not only the number of questions answered correctly but also the types of question answered (difficulty, discrimination and guessing). Two students will have different scores if one gets a right answer to a more difficult question.

D Annex - Expiration of teacher contracts

The impact of having a teacher hired with a temporary contract could depend on the duration of the contract and on how much time he has left before the end of this contract. If the contract arrives to its end soon, it could put pressure on the teacher whereas if the teacher has six years before the end of his contract he would not face such a pressure. Unfortunately, we do not have a variable indicating how much time is left before the end of the contract. Half of the contract teachers has been recruited for five years and 40% have a contract that lasts one year only or less. Very few have a contract that lasts more than five years. To construct an indicator assessing whereas the contract of the teacher expires soon, we use two variables : the duration of the contract and for how long the teacher has been employed in his current school. The latter variable is specified in brackets : for less than one year, between one and three years or for more than three years. Table D1 describes how we compute this proxy. For instance, we consider that teachers with a one-year contract face pressure as their contracts will end soon. On the contrary, teachers with a five-years contract who have been employed in their current school since less than one year are assumed not to feel such a pressure. We are aware of the limitations of such a measure as we do not know if the contract is the first they have in the current school, however our measure is the best proxy we can have.

Table D1: Construction of the proxy assessing whereas the contract of the teacher expires soon

Duration of the contract	Experience in the current school	Estimation of years left before renewal	Contract expires soon
< 1 year	No matter	< 1 year	Yes
1 year	No matter	1 year or less	Yes
2 years	< 1 year	1-2 years	No
2 years	1-3 years	1 year or less	Yes
3 years	< 1 year	2-3 years	No
3 years	1-3 years	2-3 years	No
3 years	> 3 years	1 year of less	Yes
5 years	< 1 year	4-5 years	No
5 years	1-3 years	2-4 years	No
5 years	> 3 years	2 years or less	Yes

Continued on next page

Following the previous table

Note : For teachers with a 3 years contract who have been in the school for one to three years, we assume that they had a one-year contract in the school as it is the case of many teachers.

Source : Author, using LEAPS database

E Annex - Robustness tests

E.1 Changes in score measurement

Instead of IRT subject-specific scores computed through maximum likelihood procedures, we use two different measures of scores. The first measure is a Classical Test Theory measurement where the scores are simply the standardized (by year and subject) summed of correct responses. When a child did not answer a question, it was treated as a wrong answer. The second method used is the IRT Expected A Posteriori (EAP) scores which uses a Bayesian modified curve to compute scores. The previous results remain the same (Tables E1 and E2). Local teachers are more effective especially when it comes to teaching Mathematics. Teacher experience has a low impact on students' achievement and only significant for Urdu. Non permanent contract teachers and higher paid teachers are associated with higher students' achievement in the three subjects. A teacher who has another job outside the school tends to have students with lower achievements in Urdu.

Table E1: Gain model - Classical test score measures

	(1)	(2)	(3)	(4)
Dep var : Classical gain score	English	Math	Urdu	Overall
Child underweight : BMI-for-age <-2sd	0.070 (0.058)	-0.011 (0.059)	-0.018 (0.052)	0.012 (0.049)
Child overweight : BMI-for-age >2sd	0.081 (0.154)	-0.008 (0.143)	0.094 (0.125)	0.107 (0.122)
Wealth index	0.068*** (0.021)	0.111*** (0.021)	0.110*** (0.019)	0.133*** (0.018)
Female Teacher	0.044 (0.111)	-0.225** (0.095)	-0.200** (0.100)	-0.239*** (0.084)
Same gender Teacher	-0.295*** (0.114)	-0.031 (0.091)	-0.102 (0.095)	-0.142* (0.084)
Local teacher	0.169*** (0.057)	0.172*** (0.055)	0.115** (0.050)	0.180*** (0.045)
Teacher Exp	0.021* (0.013)	0.011 (0.012)	0.041*** (0.013)	0.025** (0.011)
Teacher Exp ²	-0.001 (0.000)	-0.000 (0.000)	-0.001*** (0.000)	-0.000* (0.000)
Teacher education	0.245 (0.181)	-0.063 (0.149)	-0.058 (0.163)	0.119 (0.121)
Teacher education ²	-0.010 (0.008)	0.003 (0.006)	0.003 (0.007)	-0.004 (0.005)
Non-permanent contract	0.204** (0.100)	0.279*** (0.102)	0.427*** (0.096)	0.437*** (0.091)
PTC training	0.078 (0.072)	-0.249*** (0.073)	-0.173*** (0.065)	-0.124** (0.059)
CT training	-0.096 (0.078)	-0.171** (0.082)	-0.256*** (0.072)	-0.158** (0.067)
No training	-0.140	-0.188*	-0.212**	-0.141

Continued on next page

Following the previous table

	(0.100)	(0.110)	(0.094)	(0.086)
Log teacher monthly wage	0.228***	0.242***	0.228***	0.452***
	(0.078)	(0.074)	(0.069)	(0.062)
Bonus for pupils' performance	0.017	0.040	0.107*	0.042
	(0.069)	(0.069)	(0.064)	(0.058)
Bonus for other reasons	0.186	-0.011	0.131	-0.051
	(0.128)	(0.138)	(0.134)	(0.122)
Teacher absence	-0.019**	-0.011	-0.014**	-0.016**
	(0.008)	(0.009)	(0.007)	(0.007)
Teacher teaches outside	0.139**	0.173**	0.060	0.156***
	(0.062)	(0.068)	(0.060)	(0.057)
Teacher other work	0.091*	-0.028	-0.094*	-0.003
	(0.053)	(0.053)	(0.048)	(0.044)
Class size	-0.010***	-0.017***	-0.014***	-0.017***
	(0.002)	(0.003)	(0.003)	(0.002)
% girls in the class	0.573**	0.700***	0.110	0.519** ζ
	(0.256)	(0.258)	(0.255)	(0.227)
% with English books	-0.161			-0.411* ζ
	(0.128)			(0.244)
% with Math books		-0.374***		0.199
		(0.127)		(0.157)
% with Urdu books			-0.272**	-0.180
			(0.117)	(0.197)
% with Desks	0.166***	0.054	0.090**	0.109***
	(0.049)	(0.048)	(0.045)	(0.041)
% with Chairs	0.025	-0.002	0.018	0.095*
	(0.062)	(0.069)	(0.058)	(0.055)
% with Blackboards	-0.024	-0.050	-0.121**	-0.079
	(0.058)	(0.059)	(0.057)	(0.048)
Observations	11181	11181	11181	11181
Adjusted R ²	0.056	0.068	0.067	0.120
FE Schools	Yes	Yes	Yes	Yes
FE Teachers	No	No	No	No
FE Students	Yes	Yes	Yes	Yes

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

Teacher absence represents self-reported days of absence last month.

Reference categories : Parents have no education, the child is a boy, teacher has followed a BED training program, he cannot receive a bonus, he does not have another job and he has a permanent contract.

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

Table E2: Gain model with school and student fixed-effects- EAP IRT score measures

Dep var : IRT gain score - EAP	(1) English	(2) Math	(3) Urdu
Child underweight : BMI-for-age (WHO) <-2sd	0.022 (0.057)	0.016 (0.062)	0.017 (0.058)
Child overweight : BMI-for-age (WHO) >2sd	0.031 (0.133)	-0.015 (0.142)	0.028 (0.138)
Wealth index	0.046** (0.021)	0.098*** (0.022)	0.098*** (0.021)
Female Teacher	-0.237** (0.105)	-0.283*** (0.104)	-0.174* (0.105)
Same gender Teacher	-0.148 (0.101)	-0.000 (0.095)	-0.076 (0.095)
Local teacher	0.072 (0.056)	0.190*** (0.058)	0.055 (0.055)
Teacher Exp	0.020 (0.013)	-0.003 (0.012)	0.032** (0.014)
Teacher Exp ²	-0.001 (0.000)	0.000 (0.000)	-0.001* (0.000)
Teacher education	0.206 (0.174)	0.156 (0.191)	0.062 (0.158)
Teacher education ²	-0.008 (0.007)	-0.006 (0.008)	-0.001 (0.007)
Non-permanent contract	0.286*** (0.093)	0.181* (0.107)	0.310*** (0.095)
PTC training	0.027 (0.074)	-0.146* (0.083)	-0.120* (0.071)
CT training	-0.118 (0.079)	-0.059 (0.087)	-0.197** (0.079)
No training	-0.131 (0.103)	-0.138 (0.117)	-0.170* (0.100)
Log teacher monthly wage	0.191** (0.076)	0.279*** (0.079)	0.187** (0.073)
Bonus for pupils' performance	0.056 (0.067)	-0.010 (0.069)	0.089 (0.066)
Bonus for other reasons	0.251** (0.128)	0.111 (0.150)	0.191 (0.136)
Teacher absence	-0.008 (0.008)	-0.008 (0.008)	-0.019*** (0.007)
Teacher teaches outside	0.064 (0.059)	0.106 (0.069)	0.070 (0.063)
Teacher other work	0.076 (0.052)	-0.037 (0.058)	-0.135** (0.053)
Class size	-0.009*** (0.002)	-0.016*** (0.003)	-0.011*** (0.003)
% girls in the class	0.477* (0.248)	0.579** (0.282)	0.377 (0.292)
% with English books	-0.284** (0.130)		
% with Math books		-0.389*** (0.133)	

Continued on next page

Following the previous table

% with Urdu books			-0.365*** (0.125)
% with Desks	0.177*** (0.047)	0.039 (0.049)	0.074 (0.047)
% with Chairs	0.006 (0.058)	0.041 (0.072)	0.006 (0.059)
% with Blackboards	0.032 (0.057)	-0.075 (0.061)	-0.096 (0.061)
Observations	11181	11181	11181
Adjusted R ²	0.041	0.057	0.049
FE Schools	Yes	Yes	Yes
FE Teachers	No	No	No
FE Students	Yes	Yes	Yes

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

Teacher absence represents self-reported days of absence last month.

Reference categories : Parents have no education, the child is a boy, teacher has followed a BED training program, he cannot receive a bonus, he does not have another job and he has a permanent contract.

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

E.2 Changes in persistence rate

Following Harris and Sass (2011), we examine the robustness of our results to changes in the assumed value of the persistence rate δ . The gain model with school and student fixed effects is estimated with degrees of persistence varying parametrically from 1 (complete persistence) to 0 (no persistence) in increments of 0.2. Tables E3, E4 and E5 present the results for gains in respectively English, Mathematics and Urdu. The first columns of these tables assume no persistence of previous knowledge ($\delta = 0$) and therefore represent the estimates of the contemporaneous model (equation 6). The last columns present results when a complete persistence rate is assumed ($\delta = 1$) corresponding to the gain model (equation 7 and table 6).

Table E3: Gain model with students fixed-effects - Persistence rate changes - English

	Dep var : IRT Gain score in English					
	$\delta = 0$	$\delta = 0.2$	$\delta = 0.4$	$\delta = 0.6$	$\delta = 0.8$	$\delta = 1$
Child underweight : BMI-for-age <-2sd	0.027 (0.023)	0.028 (0.034)	0.025 (0.039)	0.021 (0.044)	0.021 (0.044)	0.015 (0.056)
Child overweight : BMI-for-age >2sd	0.126** (0.050)	0.086 (0.076)	0.071 (0.083)	0.056 (0.094)	0.056 (0.094)	0.027 (0.120)
Wealth index	0.050*** (0.008)	0.064*** (0.013)	0.063*** (0.014)	0.062*** (0.016)	0.062*** (0.016)	0.060*** (0.020)
Female Teacher	0.046 (0.041)	-0.019 (0.060)	-0.041 (0.069)	-0.063 (0.078)	-0.063 (0.078)	-0.107 (0.098)
Same gender Teacher	0.017 (0.036)	-0.097* (0.054)	-0.120* (0.062)	-0.142** (0.071)	-0.142** (0.071)	-0.187** (0.090)
Local teacher	0.064*** (0.017)	0.084** (0.034)	0.090** (0.038)	0.097** (0.043)	0.097** (0.043)	0.110** (0.053)
Teacher Exp	0.002 (0.005)	0.003 (0.008)	0.003 (0.008)	0.004 (0.009)	0.004 (0.009)	0.006 (0.012)
Teacher Exp ²	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Teacher education	-0.033** (0.015)	0.098 (0.098)	0.125 (0.111)	0.152 (0.126)	0.152 (0.126)	0.206 (0.160)
Teacher education ²	0.001*** (0.000)	-0.004 (0.004)	-0.005 (0.005)	-0.006 (0.005)	-0.006 (0.005)	-0.008 (0.007)
Non-permanent contract	0.102*** (0.031)	0.177*** (0.056)	0.200*** (0.061)	0.223*** (0.068)	0.223*** (0.068)	0.269*** (0.084)
PTC training	-0.048 (0.029)	-0.009 (0.042)	0.006 (0.048)	0.021 (0.054)	0.021 (0.054)	0.051 (0.068)
CT training	-0.104*** (0.031)	-0.098** (0.046)	-0.100* (0.052)	-0.101* (0.059)	-0.101* (0.059)	-0.105 (0.074)
No training	-0.026 (0.033)	-0.106* (0.058)	-0.114* (0.066)	-0.123 (0.075)	-0.123 (0.075)	-0.139 (0.095)
Log teacher monthly wage	0.333*** (0.029)	0.270*** (0.044)	0.268*** (0.051)	0.266*** (0.060)	0.266*** (0.060)	0.262*** (0.079)

Bonus for pupils' performance	-0.026 (0.024)	0.038 (0.038)	0.052 (0.043)	0.066 (0.049)	0.066 (0.049)	0.094 (0.061)
Bonus for other reasons	-0.006 (0.046)	0.033 (0.074)	0.089 (0.085)	0.145 (0.098)	0.145 (0.098)	0.258** (0.125)
Teacher absence	0.000 (0.003)	-0.001 (0.004)	-0.001 (0.005)	-0.001 (0.006)	-0.001 (0.006)	-0.000 (0.007)
Teacher teaches outside	-0.004 (0.022)	0.041 (0.035)	0.043 (0.040)	0.045 (0.045)	0.045 (0.045)	0.049 (0.057)
Teacher other work	0.008 (0.022)	0.042 (0.034)	0.042 (0.038)	0.043 (0.043)	0.043 (0.043)	0.044 (0.054)
Class size	-0.006*** (0.001)	-0.011*** (0.002)	-0.010*** (0.002)	-0.009*** (0.002)	-0.009*** (0.002)	-0.008*** (0.003)
% girls in the class	0.012 (0.090)	0.018 (0.141)	0.123 (0.156)	0.228 (0.175)	0.228 (0.175)	0.437** (0.221)
% with English books	0.127*** (0.038)	-0.184** (0.072)	-0.200** (0.081)	-0.217** (0.092)	-0.217** (0.092)	-0.250** (0.116)
% with Desks	0.098*** (0.020)	0.057** (0.027)	0.082*** (0.030)	0.107*** (0.035)	0.107*** (0.035)	0.157*** (0.044)
% with Chairs	0.031 (0.025)	0.052 (0.034)	0.051 (0.038)	0.049 (0.043)	0.049 (0.043)	0.046 (0.054)
% with Blackboards	0.026 (0.023)	0.038 (0.032)	0.033 (0.036)	0.029 (0.041)	0.029 (0.041)	0.020 (0.052)
Observations	18062	11181	11181	11181	11181	11181
Adjusted R ²	0.073	0.072	0.058	0.048	0.048	0.035
FE Schools	Yes	Yes	Yes	Yes	Yes	Yes
FE Teachers	No	No	No	No	No	No
FE Students	Yes	Yes	Yes	Yes	Yes	Yes

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

Reference categories : Teacher has followed a BED training program, he cannot receive a bonus, he does not have another job, he has been absent at least one day during the past month and he has a permanent contract. Control variables : Child health, household wealth index, dummies for being not promoted, % of students in the class with English, Math, Urdu books, with desks, chairs and backboards, % of children not promoted and % of girls in the class .

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

Table E4: Gain model with students fixed-effects - Persistence rate changes - Mathematics

	Dep var : IRT Gain score in Mathematics					
	$\delta = 0$	$\delta = 0.2$	$\delta = 0.4$	$\delta = 0.6$	$\delta = 0.8$	$\delta = 1$
Child underweight : BMI-for-age<-2sd	0.038 (0.024)	-0.012 (0.035)	-0.011 (0.040)	-0.010 (0.046)	-0.010 (0.046)	-0.008 (0.058)
Child overweight : BMI-for-age>2sd	0.116** (0.055)	0.069 (0.078)	0.035 (0.087)	0.001 (0.098)	0.001 (0.098)	-0.067 (0.125)
Wealth index	0.071*** (0.009)	0.081*** (0.013)	0.083*** (0.015)	0.085*** (0.017)	0.085*** (0.017)	0.089*** (0.021)
Female Teacher	-0.041 (0.043)	-0.173*** (0.065)	-0.194*** (0.072)	-0.215*** (0.079)	-0.215*** (0.079)	-0.257*** (0.097)
Same gender Teacher	0.095** (0.038)	0.019 (0.058)	-0.001 (0.064)	-0.022 (0.072)	-0.022 (0.072)	-0.063 (0.088)
Local teacher	0.119*** (0.019)	0.133*** (0.034)	0.141*** (0.038)	0.149*** (0.044)	0.149*** (0.044)	0.166*** (0.056)
Teacher Exp	0.009* (0.005)	-0.003 (0.007)	-0.002 (0.008)	-0.001 (0.009)	-0.001 (0.009)	0.000 (0.011)
Teacher Exp ²	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Teacher education	0.020 (0.017)	0.098 (0.113)	0.118 (0.126)	0.137 (0.142)	0.137 (0.142)	0.175 (0.177)
Teacher education ²	-0.001* (0.000)	-0.003 (0.005)	-0.004 (0.005)	-0.005 (0.006)	-0.005 (0.006)	-0.007 (0.008)
Non-permanent contract	0.158*** (0.037)	0.188*** (0.066)	0.192*** (0.073)	0.197** (0.080)	0.197** (0.080)	0.206** (0.097)
PTC training	-0.107*** (0.033)	-0.083* (0.046)	-0.101* (0.053)	-0.119* (0.061)	-0.119* (0.061)	-0.155* (0.079)
CT training	-0.030 (0.036)	-0.063 (0.050)	-0.076 (0.057)	-0.089 (0.066)	-0.089 (0.066)	-0.115 (0.084)
No training	-0.040 (0.039)	-0.044 (0.068)	-0.071 (0.079)	-0.097 (0.090)	-0.097 (0.090)	-0.150 (0.116)
Log teacher monthly wage	0.351*** (0.031)	0.352*** (0.047)	0.321*** (0.053)	0.289*** (0.059)	0.289*** (0.059)	0.227*** (0.073)

Bonus for pupils' performance	0.012 (0.028)	0.010 (0.039)	0.012 (0.044)	0.013 (0.050)	0.013 (0.050)	0.015 (0.063)
Bonus for other reasons	-0.079 (0.051)	-0.137* (0.080)	-0.089 (0.092)	-0.041 (0.106)	-0.041 (0.106)	0.056 (0.137)
Teacher absence	0.004 (0.003)	0.001 (0.005)	0.000 (0.005)	-0.001 (0.006)	-0.001 (0.006)	-0.003 (0.007)
Teacher teaches outside	0.034 (0.024)	0.043 (0.040)	0.052 (0.045)	0.061 (0.052)	0.061 (0.052)	0.078 (0.066)
Teacher other work	-0.004 (0.022)	0.033 (0.035)	0.026 (0.040)	0.019 (0.045)	0.019 (0.045)	0.005 (0.056)
Class size	-0.009*** (0.001)	-0.013*** (0.002)	-0.014*** (0.002)	-0.014*** (0.002)	-0.014*** (0.002)	-0.015*** (0.002)
% girls in the class	0.037 (0.106)	-0.034 (0.185)	0.053 (0.207)	0.141 (0.232)	0.141 (0.232)	0.315 (0.287)
% with Math books	-0.027 (0.038)	-0.102 (0.082)	-0.145 (0.093)	-0.187* (0.106)	-0.187* (0.106)	-0.271** (0.133)
% with Desks	0.083*** (0.021)	0.005 (0.028)	0.014 (0.031)	0.023 (0.035)	0.023 (0.035)	0.040 (0.045)
% with Chairs	0.088*** (0.027)	0.085** (0.041)	0.062 (0.047)	0.039 (0.054)	0.039 (0.054)	-0.008 (0.070)
% with Blackboards	-0.037 (0.025)	-0.061* (0.035)	-0.064 (0.039)	-0.066 (0.045)	-0.066 (0.045)	-0.072 (0.056)
Observations	18062	11181	11181	11181	11181	11181
Adjusted R ²	0.091	0.100	0.082	0.067	0.067	0.049
FE Schools	Yes	Yes	Yes	Yes	Yes	Yes
FE Teachers	No	No	No	No	No	No
FE Students	Yes	Yes	Yes	Yes	Yes	Yes

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

Reference categories : Teacher has followed a BED training program, he cannot receive a bonus, he does not have another job, he has been absent at least one day during the past month and he has a permanent contract. Control variables : Child health, household wealth index, dummies for being not promoted, % of students in the class with English, Math, Urdu books, with desks, chairs and backboards, % of children not promoted and % of girls in the class .

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

Table E5: Gain model with students fixed-effects - Persistence rate changes - Urdu

	Dep var : IRT Gain score in Mathematics					
	$\delta = 0$	$\delta = 0.2$	$\delta = 0.4$	$\delta = 0.6$	$\delta = 0.8$	$\delta = 1$
Child underweight : BMI-for-age <-2sd	0.039*	0.019	0.015	0.012	0.012	0.004
	(0.023)	(0.034)	(0.039)	(0.045)	(0.045)	(0.057)
Child overweight : BMI-for-age >2sd	0.089*	0.102	0.092	0.082	0.082	0.063
	(0.050)	(0.085)	(0.095)	(0.106)	(0.106)	(0.131)
Wealth index	0.069***	0.086***	0.089***	0.092***	0.092***	0.098***
	(0.009)	(0.013)	(0.015)	(0.017)	(0.017)	(0.021)
Female Teacher	-0.061	-0.136**	-0.150**	-0.164**	-0.164**	-0.192*
	(0.044)	(0.068)	(0.074)	(0.081)	(0.081)	(0.098)
Same gender Teacher	0.075*	-0.035	-0.048	-0.061	-0.061	-0.088
	(0.039)	(0.063)	(0.068)	(0.074)	(0.074)	(0.088)
Local teacher	0.103***	0.051	0.054	0.058	0.058	0.064
	(0.018)	(0.033)	(0.037)	(0.042)	(0.042)	(0.053)
Teacher Exp	0.009*	0.015**	0.019**	0.022**	0.022**	0.029**
	(0.005)	(0.007)	(0.009)	(0.010)	(0.010)	(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	0.038**	0.087	0.108	0.130	0.130	0.172
	(0.016)	(0.097)	(0.107)	(0.120)	(0.120)	(0.151)
Teacher education ²	-0.001**	-0.003	-0.004	-0.005	-0.005	-0.007
	(0.000)	(0.004)	(0.005)	(0.005)	(0.005)	(0.006)
Non-permanent contract	0.103***	0.226***	0.237***	0.247***	0.247***	0.269***
	(0.033)	(0.059)	(0.065)	(0.071)	(0.071)	(0.087)
PTC training	-0.078**	-0.071	-0.084*	-0.097*	-0.097*	-0.124*
	(0.032)	(0.045)	(0.051)	(0.058)	(0.058)	(0.073)
CT training	-0.052	-0.096*	-0.123**	-0.151**	-0.151**	-0.206**
	(0.033)	(0.052)	(0.059)	(0.067)	(0.067)	(0.085)
No training	-0.050	-0.051	-0.078	-0.106	-0.106	-0.162*
	(0.037)	(0.062)	(0.070)	(0.078)	(0.078)	(0.097)
Log teacher monthly wage	0.303***	0.262***	0.234***	0.207***	0.207***	0.151**
	(0.029)	(0.044)	(0.050)	(0.057)	(0.057)	(0.073)

Bonus for pupils' performance	0.022 (0.026)	0.061* (0.037)	0.075* (0.041)	0.090* (0.047)	0.090* (0.047)	0.118** (0.059)
Bonus for other reasons	0.001 (0.050)	-0.025 (0.077)	0.033 (0.086)	0.091 (0.097)	0.091 (0.097)	0.207* (0.120)
Teacher absence	0.001 (0.003)	-0.004 (0.004)	-0.006 (0.005)	-0.009* (0.005)	-0.009* (0.005)	-0.013** (0.006)
Teacher teaches outside	0.007 (0.022)	0.058 (0.036)	0.064 (0.041)	0.069 (0.046)	0.069 (0.046)	0.080 (0.058)
Teacher other work	-0.035 (0.022)	-0.054 (0.034)	-0.070* (0.038)	-0.086** (0.043)	-0.086** (0.043)	-0.118** (0.054)
Class size	-0.007*** (0.001)	-0.011*** (0.002)	-0.011*** (0.002)	-0.011*** (0.002)	-0.011*** (0.002)	-0.010*** (0.002)
% with Urdu books	0.023 (0.036)	-0.257*** (0.073)	-0.301*** (0.083)	-0.345*** (0.094)	-0.345*** (0.094)	-0.433*** (0.120)
% with Desks	0.057*** (0.020)	-0.001 (0.026)	0.012 (0.030)	0.026 (0.033)	0.026 (0.033)	0.054 (0.042)
% with Chairs	0.106*** (0.023)	0.073** (0.033)	0.055 (0.037)	0.037 (0.042)	0.037 (0.042)	0.001 (0.055)
% with Blackboards	0.006 (0.025)	-0.003 (0.035)	-0.018 (0.040)	-0.033 (0.045)	-0.033 (0.045)	-0.063 (0.056)
% girls in the class	0.047 (0.113)	-0.005 (0.196)	0.109 (0.213)	0.224 (0.234)	0.224 (0.234)	0.454 (0.282)
Observations	18062	11181	11181	11181	11181	11181
Adjusted R ²	0.079	0.085	0.071	0.060	0.060	0.048
FE Schools	Yes	Yes	Yes	Yes	Yes	Yes
FE Teachers	No	No	No	No	No	No
FE Students	Yes	Yes	Yes	Yes	Yes	Yes

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

Reference categories : Teacher has followed a BED training program, he cannot receive a bonus, he does not have another job, he has been absent at least one day during the past month and he has a permanent contract. Control variables : Child health, household wealth index, dummies for being not promoted, % of students in the class with English, Math, Urdu books, with desks, chairs and backboards, % of children not promoted and % of girls in the class .

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

The negative effect of female teachers on students' achievement in Mathematics and Urdu is smaller when lower persistence is assumed and it is no longer significant in the case of the contemporaneous model. Having a local teacher benefits students' achievement in English and Mathematics whatever the persistence rate assumed but the effect is larger when persistence is higher. The positive effect of non permanent contract teacher on English, Mathematics and Urdu achievement holds for all persistence rates but the magnitude of this impact decreases slightly when persistence diminishes. Teacher wages significantly affect students' achievement in all three subjects no matter the persistence rates assumed. However, it is interesting to note that the magnitude of this effect tends to diminish when persistence rates increase. Therefore, if students' achievement are highly dependent on prior knowledge, having a non permanent contract will be associated with higher gains in achievement than having a better paid teacher whereas it is the contrary if prior knowledge tends to decay rapidly.

Other previous findings are more dependent on the degree of persistence assumed. The negative effect of same-gender teacher upon English achievement is not significant at 5% when a persistence rate lower than 0.6 is assumed. The negative impact of being taught by a teacher who has another job is only significant on Urdu achievement if persistence rates superior to 0.2 are assumed. The magnitude of this impact increases with higher persistence rate.

E.3 Changes in sub-samples

The characteristics affecting skills may vary according to the child's gender. Analyzing separately boys and girls is even more important in Pakistan since the Pakistani educational system is segregated between girls and boys (Westbrook et al., 2009). Once again, the main variables - teachers' salary and his contract - remain relevant for both gender (Table E6). Recruiting more local teachers could reduce the gender gap in academic achievement as it impacts girls more than boys. Having a same-gender teacher is associated with significant higher scores for girls than for boys. Older training programs (PTC and CT) are particularly detrimental to girls. When girls are in a classroom with other girls it is associated with higher gains in English and Urdu scores. Peer effects due to gender composition of the class are important factors explaining girls' achievement.

Table E6: Gain model with school and student fixed-effects : gender

Dep var : IRT gain score - ML	Girls			Boys		
	English	Math	Urdu	English	Math	Urdu
Child underweight : BMI-for-age <-2sd	0.088 (0.075)	-0.045 (0.086)	0.016 (0.080)	-0.035 (0.081)	0.031 (0.078)	0.005 (0.081)
Child overweight : BMI-for-age >2sd	0.394* (0.208)	0.639*** (0.232)	0.769*** (0.253)	-0.099 (0.144)	-0.307** (0.141)	-0.187 (0.144)
Wealth index	0.056** (0.028)	0.105*** (0.033)	0.096*** (0.030)	0.065** (0.029)	0.068** (0.028)	0.095*** (0.030)
Female Teacher	-0.260* (0.142)	-0.220 (0.140)	-0.162 (0.148)	0.110 (0.142)	-0.219 (0.138)	-0.193 (0.127)
Local teacher	0.207*** (0.074)	0.189** (0.077)	0.194*** (0.073)	-0.025 (0.076)	0.169** (0.082)	-0.063 (0.077)
Teacher Exp	-0.002 (0.019)	-0.015 (0.021)	0.068*** (0.021)	0.005 (0.015)	-0.001 (0.014)	0.009 (0.017)
Teacher Exp ²	0.000 (0.001)	0.001 (0.001)	-0.001* (0.001)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Teacher education	-0.131 (0.319)	0.076 (0.299)	0.139 (0.243)	0.410** (0.192)	0.291 (0.202)	0.275 (0.209)
Teacher education ²	0.006 (0.014)	-0.005 (0.013)	-0.006 (0.010)	-0.016** (0.008)	-0.010 (0.009)	-0.011 (0.009)
Non-permanent contract	0.292** (0.116)	0.058 (0.151)	0.308** (0.126)	0.286** (0.123)	0.379*** (0.124)	0.308** (0.123)
PTC training	-0.171* (0.100)	-0.370*** (0.119)	-0.318*** (0.098)	0.281*** (0.092)	-0.005 (0.110)	0.022 (0.108)
CT training	-0.203* (0.104)	-0.218* (0.131)	-0.397*** (0.126)	0.006 (0.105)	-0.081 (0.117)	-0.095 (0.117)
No training	-0.007 (0.134)	-0.034 (0.156)	-0.144 (0.121)	-0.128 (0.134)	-0.244 (0.172)	-0.112 (0.151)
Log teacher monthly wage	0.371*** (0.124)	0.303*** (0.116)	0.113 (0.111)	0.266*** (0.099)	0.247** (0.099)	0.232** (0.097)
Bonus for pupils' performance	0.184** (0.088)	0.156 (0.101)	0.134 (0.086)	0.050 (0.082)	-0.072 (0.080)	0.126 (0.083)

Bonus for other reasons	0.024 (0.193)	-0.165 (0.250)	0.102 (0.185)	0.432*** (0.159)	0.239 (0.159)	0.320** (0.157)
Teacher absence last month:0 days	-0.039 (0.052)	-0.020 (0.060)	-0.066 (0.053)	-0.141** (0.061)	-0.065 (0.060)	-0.130** (0.056)
Teacher teaches outside	0.130* (0.074)	0.180* (0.097)	0.123 (0.086)	-0.003 (0.086)	0.016 (0.094)	0.066 (0.080)
Teacher other work	0.198** (0.091)	-0.039 (0.110)	-0.211** (0.095)	-0.023 (0.066)	0.015 (0.064)	-0.101 (0.065)
Class size	-0.000 (0.003)	-0.014*** (0.003)	-0.006* (0.003)	-0.018*** (0.004)	-0.018*** (0.004)	-0.014*** (0.003)
% girls in the class	0.828*** (0.283)	0.033 (0.446)	0.794* (0.476)	0.228 (0.337)	0.819** (0.344)	0.317 (0.314)
% with English books	0.020 (0.155)			-0.470*** (0.178)		
% with Math books		0.016 (0.183)			-0.545*** (0.195)	
% with Urdu books			-0.275* (0.161)			-0.573*** (0.183)
% with Desks	-0.010 (0.062)	-0.017 (0.065)	-0.038 (0.063)	0.270*** (0.061)	0.104* (0.061)	0.129** (0.057)
% with Chairs	0.112 (0.072)	-0.044 (0.097)	-0.124 (0.080)	-0.060 (0.081)	0.011 (0.103)	0.093 (0.077)
% with Blackboards	0.040 (0.079)	-0.186** (0.083)	-0.130 (0.095)	0.048 (0.070)	0.062 (0.075)	0.029 (0.066)
Observations	5089	5089	5089	6092	6092	6092
Adjusted R ²	0.056	0.078	0.073	0.050	0.055	0.050
FE Schools	Yes	Yes	Yes	Yes	Yes	Yes
FE Teachers	No	No	No	No	No	No
FE Students	Yes	Yes	Yes	Yes	Yes	Yes

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

Reference categories : Teacher has followed a BED training program, he cannot receive a bonus, he does not have another job, he has been absent at least one day during the past month and he has a permanent contract.

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

Private and public education providers differ in many ways. Table E7 presents the descriptive statistics in these two types of schools. Pupils in public schools are different from those attending private schools. Students attending private schools perform significantly better than pupils in public schools in all subjects. They come from richer and more educated families in comparison with students in public schools. Teachers in public and private schools are drawn from two completely different groups as underlined by Andrabi et al. (2007). Teachers in private schools are predominantly women and they are less trained and experienced compared with teachers in governmental institutions. This should theoretically have a positive impact for public school but the effect is undermined by the inefficiency of training programs underlined previously. Very few teachers in private schools have a permanent contract. Teachers' wages in private schools are five times less than the average wage of a teacher in a governmental school. However, this gap could be just driven by characteristics : teachers in public schools could be paid more only because they are better trained and more educated. To control for these differences and compare teachers with similar profiles, salaries in both sectors are regressed on teacher education, experience and training. We find that teachers with identical profiles still earn 2-3 times more in government schools than their colleagues in private schools. These results are consistent with Andrabi et al. (2007). In private schools, teachers are more likely to receive bonuses. Government teachers seem to exert less effort than their private counterparts as head teachers on average report higher absenteeism. This is consistent with Andrabi et al. (2007) who that the higher rate of absenteeism in public schools is partly linked with teachers additional responsibilities.

Table E7: Descriptive statistics : public and private schools

	Public		Private		Difference	
	Mean	Sd	Mean	Sd	Diff	T-statistic
<i>Scores</i>						
Std English score	-0.25	0.9	0.64	0.9	-0.894***	(-96.76)
Std Math score	-0.15	1.0	0.37	0.9	-0.517***	(-52.70)
Std Urdu score	-0.18	0.9	0.46	1.0	-0.635***	(-65.64)
Std total score	-0.21	0.9	0.54	0.9	-0.745***	(-78.33)
<i>Child characteristics</i>						
Child age	10.17	1.7	10.01	1.7	0.160***	(9.36)
Girl	0.45	0.5	0.44	0.5	0.013**	(2.63)
Child underweight : BMI-for-age<-2sd	0.14	0.3	0.13	0.3	0.004	(1.02)
Child overweight : BMI-for-age>-2sd	0.02	0.2	0.02	0.1	0.004	(1.93)
<i>Household characteristics</i>						
Dad uneducated	0.42	0.5	0.26	0.4	0.161***	(28.61)
Dad < primary	0.09	0.3	0.08	0.3	0.017***	(4.95)
Dad primary-hig sec	0.44	0.5	0.54	0.5	-0.105***	(-17.85)
Dad > high sec	0.04	0.2	0.12	0.3	-0.073***	(-24.78)
Mum uneducated	0.71	0.5	0.52	0.5	0.187***	(33.78)
Continued on next page						

Following the previous table

Mum < primary	0.09	0.3	0.09	0.3	-0.003	(-1.00)
Mum primary-high sec	0.20	0.4	0.36	0.5	-0.161***	(-32.12)
Mum > high sec	0.01	0.1	0.03	0.2	-0.022***	(-16.22)
Elder siblings	6.68	2.5	6.94	2.0	-0.259***	(-8.83)
Wealth index	-0.31	1.3	0.63	1.5	-0.940***	(-53.83)
<i>School characteristics</i>						
School has a library	0.27	0.4	0.44	0.5	-0.168***	(-36.17)
School has computers facilities	0.01	0.1	0.33	0.5	-0.320***	(-120.85)
School has sports facilities	0.16	0.4	0.40	0.5	-0.237***	(-57.85)
School has an activity room	0.10	0.3	0.18	0.4	-0.082***	(-25.00)
School has four walls	0.70	0.5	0.97	0.2	-0.278***	(-68.95)
School has fans/rooms coolers	0.68	0.5	0.96	0.2	-0.273***	(-66.24)
School has electricity	0.75	0.4	0.98	0.1	-0.227***	(-60.32)
Admission fee (Grades 1 to 3)	1.43	14.4	125.53	185.1	-124.1***	(-124.64)
Annual school fee (Grades 1 to 3)	20.38	165.3	1292.32	934.5	-1271.9***	(-245.29)
Admission fee (Grades 4 to 5)	1.84	15.9	146.80	257.9	-145.0***	(-104.74)
Annual school fee (Grades 4 to 5)	52.49	399.5	1508.95	1062.9	-1456.5***	(-219.85)
<i>Teacher characteristics</i>						
Female Teacher	0.45	0.5	0.79	0.4	-0.342***	(-71.27)
Same gender Teacher	0.96	0.2	0.50	0.5	0.455***	(142.63)
Teacher Experience	13.93	7.7	3.98	4.4	9.944***	(136.72)
Teacher training	0.97	0.2	0.25	0.4	0.716***	(257.69)
Teacher education	11.54	1.7	11.58	1.3	-0.043*	(-2.55)
Matric	0.41	0.5	0.34	0.5	0.0703***	(14.23)
FA-FSc	0.24	0.4	0.43	0.5	-0.198***	(-43.81)
BA-BSc	0.26	0.4	0.21	0.4	0.0544***	(12.48)
MA or higher	0.09	0.3	0.02	0.1	0.068***	(26.57)
Permanent contract	0.83	0.4	0.02	0.1	0.809***	(207.75)
Temporary contract	0.17	0.4	0.87	0.3	-0.700***	(-162.79)
Monthly teaching wage (Rs)	6519.57	1988.8	1311.31	1014.4	5208.3***	(289.64)
Can receive bonus	0.26	0.4	0.43	0.5	-0.166***	(-34.68)
Did receive bonus	0.03	0.2	0.19	0.4	-0.160***	(-59.34)
Teacher works in agriculture	0.19	0.4	0.07	0.2	0.125***	(33.27)
Teacher runs a business	0.03	0.2	0.03	0.2	-0.001	(-0.56)
Teacher teaches outside school	0.04	0.2	0.33	0.5	-0.296***	(-93.44)
Teacher has another work	0.02	0.1	0.08	0.3	-0.061***	(-29.99)
Teacher absence	2.22	2.9	1.34	2.3	0.883***	(31.70)
Observations	35092		13654		48746	

Notes : t statistics in parentheses : * $p < .05$, ** $p < .01$, *** $p < .001$

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

As there may be different factors influencing the efficiency of public schools and private schools, we run separate regressions, using the gain model, for public and private schools. When we analyze the inputs influencing students' achievement, we observe significant differences between private and public institutions (Table E8). First, the positive gain in achievement associated with contract teachers holds for both types of schools even if it is slightly higher in private schools. Female teachers are less effective in private schools than male teachers. This result is consistent with the specific context of Pakistan, where private schools have been recruiting untrained, less educated women and

pay them less than male teachers in order to keep their costs low, charge low fees to attract children from poor households (Andrabi et al., 2008). Therefore, women in private schools probably feel less motivated in their job. Wages and bonus policies are relevant tools to motivate teachers to perform better especially in public schools.

Table E8: Gain model with students fixed-effects : private and public schools

Dep var : IRT gain score - ML	Private School			Public Schools		
	English	Math	Urdu	English	Math	Urdu
Child underweight : BMI-for-age <-2sd	0.010 (0.082)	0.121 (0.088)	0.059 (0.086)	-0.002 (0.076)	-0.089 (0.077)	-0.021 (0.078)
Child overweight : BMI-for-age >2sd	0.093 (0.188)	-0.036 (0.222)	0.145 (0.229)	0.016 (0.151)	-0.048 (0.151)	0.063 (0.159)
Wealth index	0.101*** (0.028)	0.062** (0.031)	0.064** (0.028)	0.039 (0.028)	0.083*** (0.029)	0.108*** (0.029)
Female Teacher	-0.176 (0.112)	-0.429*** (0.110)	-0.290*** (0.100)	-0.318 (0.311)	-0.521 (0.387)	0.094 (0.533)
Same gender Teacher	-0.136 (0.090)	-0.031 (0.085)	-0.122 (0.078)	-0.609* (0.341)	-0.784** (0.379)	-0.079 (0.529)
Local teacher	0.135** (0.067)	0.110 (0.076)	-0.043 (0.066)	0.096 (0.082)	0.182** (0.084)	0.147* (0.084)
Teacher Exp	-0.009 (0.018)	-0.026 (0.019)	0.051*** (0.019)	0.019 (0.016)	0.004 (0.014)	0.018 (0.018)
Teacher Exp ²	0.000 (0.000)	0.000 (0.000)	-0.001*** (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.001)
Teacher education	0.199 (0.168)	0.209 (0.206)	0.171 (0.157)	0.057 (0.372)	-0.085 (0.414)	0.091 (0.424)
Teacher education ²	-0.008 (0.007)	-0.007 (0.009)	-0.009 (0.007)	0.000 (0.016)	0.004 (0.018)	-0.001 (0.018)
Non-permanent contract	0.361*** (0.114)	0.254** (0.129)	0.273** (0.120)	0.332** (0.137)	0.170 (0.142)	0.274** (0.136)
PTC training	0.018 (0.094)	-0.299** (0.137)	-0.106 (0.097)	0.121 (0.102)	-0.188* (0.110)	-0.071 (0.114)
CT training	-0.064 (0.129)	-0.056 (0.206)	-0.054 (0.151)	-0.147 (0.092)	-0.174* (0.095)	-0.204** (0.103)
No training	-0.397*** (0.120)	-0.455** (0.194)	-0.104 (0.124)	0.642** (0.301)	0.150 (0.256)	-0.102 (0.250)
Log teacher monthly wage	0.041 (0.093)	0.181* (0.103)	0.062 (0.089)	0.510*** (0.136)	0.285** (0.114)	0.219* (0.120)

Bonus for pupils' performance	-0.051 (0.072)	-0.033 (0.077)	0.013 (0.068)	0.185 (0.118)	0.051 (0.118)	0.270** (0.118)
Bonus for other reasons	0.073 (0.147)	-0.420** (0.164)	0.038 (0.139)	0.460** (0.233)	0.560** (0.232)	0.500** (0.234)
Teacher absence	-0.031** (0.015)	-0.043** (0.018)	-0.017 (0.015)	0.007 (0.009)	0.013 (0.008)	-0.011 (0.007)
Teacher teaches outside	0.095 (0.060)	0.022 (0.073)	0.013 (0.060)	-0.087 (0.185)	0.050 (0.188)	0.250 (0.189)
Teacher other work	0.047 (0.090)	0.014 (0.105)	-0.163* (0.084)	0.021 (0.067)	0.002 (0.067)	-0.103 (0.068)
Class size	-0.006 (0.005)	0.004 (0.005)	0.000 (0.005)	-0.007** (0.003)	-0.020*** (0.003)	-0.012*** (0.003)
% girls in the class	0.359 (0.283)	0.437 (0.305)	0.510* (0.281)	0.797* (0.413)	-0.040 (0.521)	0.355 (0.561)
% with English books	-0.129 (0.165)			-0.359** (0.172)		
% with Math books		-0.547*** (0.209)			-0.367** (0.187)	
% with Urdu books			-0.504** (0.196)			-0.531*** (0.163)
% with Desks	0.048 (0.075)	0.169* (0.093)	0.192** (0.078)	0.201*** (0.056)	-0.004 (0.053)	-0.006 (0.051)
% with Chairs	0.064 (0.068)	0.228*** (0.088)	0.078 (0.074)	-0.046 (0.098)	-0.267** (0.120)	-0.027 (0.090)
% with Blackboards	0.282*** (0.101)	0.466*** (0.146)	0.176 (0.116)	-0.023 (0.060)	-0.148** (0.061)	-0.111* (0.064)
Observations	3635	3635	3635	7514	7514	7514
Adjusted R ²	0.071	0.083	0.056	0.032	0.051	0.047
FE Schools	Yes	Yes	Yes	Yes	Yes	Yes
FE Teachers	No	No	No	No	No	No
FE Students	Yes	Yes	Yes	Yes	Yes	Yes

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

Reference categories : Teacher has followed a BED training program, he cannot receive a bonus, he does not have another job, he has been absent at least one day during the past month and he has a permanent contract.

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

E.4 Balanced Samples

In addition with these sub-sample breakdowns, to deal with attrition, estimates for the gain model with student and school fixed effects using the balanced students panel sample are presented in table E9. The previous results are robust to this sample change : having a local teacher with a temporary contract and having a better paid teacher are associated with higher gains in scores.

Table E9: Gain model with school and student fixed effects - balanced sample

Dep var : IRT gain score - ML	(1)	(2)	(3)
	English	Math	Urdu
Child underweight : BMI-for-age <-2sd	0.015 (0.056)	-0.008 (0.058)	0.004 (0.057)
Child overweight : BMI-for-age >2sd	0.027 (0.120)	-0.067 (0.125)	0.063 (0.131)
Wealth index	0.060*** (0.020)	0.089*** (0.021)	0.098*** (0.021)
Female Teacher	-0.107 (0.098)	-0.257*** (0.097)	-0.192* (0.098)
Same gender Teacher	-0.187** (0.090)	-0.063 (0.088)	-0.088 (0.089)
Local teacher	0.110** (0.053)	0.166*** (0.056)	0.064 (0.053)
Teacher Exp	0.006 (0.012)	0.000 (0.011)	0.029** (0.013)
Teacher Exp ²	-0.000 (0.000)	-0.000 (0.000)	-0.001 (0.000)
Teacher education	0.206 (0.160)	0.175 (0.178)	0.172 (0.151)
Teacher education ²	-0.008 (0.007)	-0.007 (0.008)	-0.007 (0.006)
Non-permanent contract	0.269*** (0.084)	0.206** (0.097)	0.269*** (0.088)
PTC training	0.051 (0.068)	-0.155* (0.079)	-0.124* (0.073)
CT training	-0.105 (0.074)	-0.115 (0.084)	-0.206** (0.085)
No training	-0.139 (0.096)	-0.150 (0.116)	-0.162* (0.097)
Log teacher monthly wage	0.262*** (0.079)	0.227*** (0.073)	0.151** (0.073)
Bonus for pupils' performance	0.094 (0.061)	0.015 (0.063)	0.118** (0.060)
Bonus for other reasons	0.258** (0.125)	0.056 (0.137)	0.207* (0.120)
Teacher absence	-0.000 (0.007)	-0.003 (0.007)	-0.013** (0.006)
Teacher teaches outside	0.049 (0.058)	0.078 (0.066)	0.080 (0.058)
Teacher other work	0.044 (0.054)	0.005 (0.056)	-0.118** (0.054)
Class size	-0.008***	-0.015***	-0.010***

Continued on next page

Following the previous table

	(0.003)	(0.002)	(0.002)
% girls in the class	0.437** (0.221)	0.315 (0.288)	0.454 (0.283)
% with English books	-0.250** (0.116)		
% with Math books		-0.271** (0.133)	
% with Urdu books			-0.433*** (0.120)
% with Desks	0.157*** (0.044)	0.040 (0.045)	0.054 (0.042)
% with Chairs	0.046 (0.054)	-0.008 (0.070)	0.001 (0.055)
% with Blackboards	0.020 (0.052)	-0.072 (0.056)	-0.063 (0.056)
Observations	8968	8968	8968
Adjusted R ²	0.032	0.047	0.045
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 sample in this regression contains only the children and the teachers surveyed three times.

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

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

As underlined previously, teacher attrition may also be an issue. Table E10 presents the results with only students and teachers surveyed three times. Having a contract teacher and having a better paid teacher are still associated with higher achievement.

Table E10: Gain model with school and student fixed effects - balanced student and teacher sample

Dep var : IRT gain score - ML	(1) English	(2) Math	(3) Urdu
Child underweight : BMI-for-age <-2sd	0.137 (0.094)	-0.088 (0.088)	-0.000 (0.089)
Child overweight : BMI-for-age >2sd	-0.297 (0.184)	-0.312 (0.208)	-0.135 (0.218)
Wealth index	0.038 (0.031)	0.034 (0.034)	0.067* (0.034)
Female Teacher	0.909 (0.783)	-3.432*** (0.860)	-2.545*** (0.800)
Same gender Teacher	-1.896*** (0.449)	-1.586*** (0.427)	-1.304*** (0.398)
Local teacher	0.219 (0.138)	0.216 (0.162)	-0.068 (0.155)
Teacher Exp	0.020 (0.023)	0.016 (0.021)	0.072*** (0.027)
Teacher Exp ²	-0.001 (0.001)	-0.000 (0.001)	-0.001* (0.001)

Continued on next page

Following the previous table

Teacher education	1.202** (0.519)	-1.380** (0.541)	-1.841*** (0.507)
Non-permanent contract	0.355 (0.385)	0.895* (0.479)	-0.125 (0.457)
PTC training	0.250 (0.159)	-0.296 (0.191)	-0.091 (0.211)
CT training	-2.831*** (1.048)	4.330*** (1.097)	0.211 (1.003)
Log teacher monthly wage	0.392** (0.186)	0.749*** (0.183)	0.702*** (0.160)
Bonus for pupils' performance	0.213 (0.254)	-0.170 (0.316)	0.383 (0.235)
Bonus for other reasons	-0.121 (0.356)	0.843* (0.430)	0.605* (0.364)
Teacher absence	-0.010 (0.010)	-0.008 (0.009)	-0.019** (0.008)
Teacher teaches outside	0.245* (0.127)	0.626*** (0.191)	0.040 (0.142)
Teacher other work	-0.076 (0.112)	-0.069 (0.109)	-0.114 (0.103)
Class size	-0.016*** (0.005)	-0.018*** (0.005)	-0.012*** (0.005)
% girls in the class	1.745*** (0.653)	0.971 (0.807)	0.029 (0.713)
% with English books	-0.294 (0.219)		
% with Math books		-0.067 (0.215)	
% with Urdu books			-0.229 (0.180)
% with Desks	0.322*** (0.074)	0.064 (0.070)	-0.044 (0.064)
% with Chairs	0.005 (0.121)	0.093 (0.174)	-0.036 (0.139)
% with Blackboards	0.069 (0.077)	0.069 (0.077)	0.003 (0.080)
Observations	3179	3179	3179
Adjusted R ²	0.060	0.075	0.071
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 sample in this regression contains only the children and the teachers surveyed three times.

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

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