Online Appendix
Teacher Value-Added in a Low-Income Country

Natalie Bau and Jishnu Das

November 5, 2019

Appendix A describes how our test score data was collected and how the tests were scored, and Appendix B details the derivation of the bias term in the estimate of the variance of the classroom effects. Appendix C documents how data entry errors in teacher ids could lead to greater bias in TVA estimates that control for child fixed effects, while Appendix D describes the details of the teacher hiring policy simulation.

Appendix A: Test Data

Data Collection

In each round of the LEAPS data collection, we tested students in math, Urdu (the vernacular), and English. To avoid the possibility of cheating, project staff, with clear instructions not to interfere, administered the test directly to students. Test booklets were retrieved after class, so there was no missing testing material. Tests were scored and equated across the four rounds using Item Response Theory, yielding scores in each subject with a mean of 0 and a standard deviation of 1 (Das and Zajonc, 2010). Item response theory weights questions differently according to their difficulty and allows us to equate tests over years so that a standard deviation gain in year 1 is equivalent to a standard deviation gain in year 4 in terms of student knowledge. The tests could be equated because we included linking questions across any two years and for some questions, across multiple years.

Table 1 provides more information on the sources of variation for the TVA calculations. In year one, since only 3rd graders were tested, very few students were observed in schools where more than one classroom was tested. In future years, some students were held back, others were promoted, and another sample of 3rd graders was added in year 3, allowing students in a larger number of classrooms to be tested. Columns 1 and 2 describe the sample used to calculate the cross-school TVA estimates. Columns 3 and 4 describe the variation used to calculate the within-school TVA measures.

Scoring Tests

To score the tests that we administered to teachers and students, we separately estimate the IRT scores for years 1 to 4 for the students, recovering an ability parameter for each student-year combination, as well as 3 parameters (per test question) that measure the difficulty of each test item. We then use the item parameters from this student-level IRT estimation as inputs into the ability parameter estimates for the

teacher. This procedure is only valid if the structural assumptions underlying this extrapolation are true. That is, although teachers can, on average, have a higher ability parameter, their likelihood of correctly answering a question conditional on ability should be fully determined by the structural 3-parameter logistic equation with the item parameters as inputs. To test this, we can use a χ^2 test comparing the observed response pattern among teachers with the predicted response pattern from the student-level IRT estimation.

In Appendix Figures A11-A13, we produce the item-level observed responses from teachers plotted against the predicted responses based on the student-level estimation separately for English, mathematics, and Urdu. Since the χ^2 test tends to be underpowered in small samples (and is never rejected in our case), we view these figures as more informative about the fit of the IRT model. The figures that we present here treat all the teachers as a single cross-section, including teachers who are tested multiple times. We also perform a similar computation retaining a single observation per teacher and find exactly the same patterns. The figures for all 3 subjects consistently show: (a) that teachers are performing at the upper range of the student ability distribution and that (b) that for most questions, the predicted response pattern from the IRT model match the patterns observed in the data. This close match could be caused by the fact that there is not much variation in the structural item curve in the upper range of the ability distribution. However, where there there is considerable variation (see, for instance, math items 38 and 40, Urdu items 13, 25, and 46 or English items 41, 44, and 48), we find similar concordance between predicted and observed responses. There are a few specific questions (for instance, math items 33 and 34, Urdu item 22, and English item 50) for which the fit is poor, but we could find no clear pattern to explain this relatively poorer fit. In math, the items were a harder division problem and a simple arithmetic problem (the teachers are less likely to get it right, conditional on ability); for Urdu, item 22 is an antonym, and for English, item 50 is a reading comprehension question where students read a passage and answer a question based on that passage. One possible explanation is that teaching has changed over time to stress these kinds of tasks more. This may lead older teachers to perform worse on these items.

Appendix B: Derivation of Bias

In this appendix, we derive an estimate of ϕ, the term capturing the sampling bias in the expression $Var(\widehat{\delta_{jst}}) = \sigma^2_{jst} + \sigma^2_{js} + \sigma^2_s + \phi$. The notation here is defined in the same way as in the body of the paper. To derive ϕ, we first note that $\widehat{\delta_{jst}} = \delta_{jst} + \frac{\sum_{i=1}^{N_{jt}} \upsilon_{ijt}}{N_{jt}}$, where $\delta_{jst} = \theta_s + \theta_{js} + \theta_{jst}$ and N_{jt} is the number of students taught by teacher j in year t. Furthermore, assume υ_{ijt} is homoskedastic with variance σ^2_υ. Then,

$$Var(\widehat{\delta_{jst}}) = Var(\delta_{jst} + \frac{\sum_{i=1}^{N_{jt}} \upsilon_{ijt}}{N_{jt}}, \delta_{jst} + \frac{\sum_{i=1}^{N_{jt}} \upsilon_{ijt}}{N_{jt}}),$$

and a little algebra shows that

$$Var(\widehat{\delta_{jst}}) = Var(\delta_{jst}) + 2E(\frac{\delta_{jst} \sum_{i}^{N_{jt}} \upsilon_{jt}}{N^2_{jt}}) + E(\frac{\sum_{i}^{N_{jt}} \upsilon_{ijt}}{N_{jt}} \times \frac{\sum_{i}^{N_{jt}} \upsilon_{ijt}}{N_{jt}}).$$

Then, we can recognize that $2E(\frac{\delta_{jst} \sum_{i}^{N_{jt}} \upsilon_{ijt}}{N^2_{jt}}) = 0$ since δ_{jst} and υ_{ijt} are independent and $E(\upsilon_{ijt}) = 0$ by construction. Additionally, we can recognize that $E(\frac{\sum_{i}^{N_{jt}} \upsilon_{ijt}}{N_{jt}} \times \frac{\sum_{i}^{N_{jt}} \upsilon_{ijt}}{N_{jt}}) = E(\frac{\sigma^2_\upsilon}{N_{jt}}) = \phi$. Then,

$$Var(\widehat{\delta_{jst}}) = Var(\delta_{js}) + \phi,$$

which is equivalent to

$$Var(\widehat{\delta_{jst}}) = \sigma^2_s + \sigma^2_{js} + \sigma^2_{jst} + \phi.$$

Appendix C: Incorrect Variation in Teacher Switching Due to Data Entry Errors

In this appendix, we show how a small amount of data misentry can lead to a large amount of bias when we include child fixed effects in the TVA estimation. Suppose that 1 percent of teacher IDs are randomly entered incorrectly. If 10 percent of students change teachers each year, when identifying variation comes only from the test scores of students who change teachers, these incorrect entries account for 9 percent of the variation. To arrive at this number, note that there are three cases where a student-year observation will provide identifying variation in a specification that includes child fixed effects: (1) the teacher ID was incorrectly entered, but no switch actually occurred (probability $= 0.01 \times 0.9 = 0.009$), (2) the teacher ID was correctly entered and a switch occurred (probability $= 0.99 \times 0.1 = .099$), and (3) the ID was incorrectly entered and a switch occurred (probability $= 0.1 \times 0.01 = 0.001$). Then the probability that the teacher ID is mis-attributed in an observation that provides identifying variation is $\frac{0.01}{(0.009+0.099+0.001)} = 0.09$.

In order to assess potential bias more formally, consider a case where students are identical and TVA is randomly distributed, so there is no correlation between a student's future teacher's TVA and his current teacher's TVA as long as she changes teachers. Now, also assume that a student has a probability p of changing teachers each year, and an ID has a probability e of being incorrectly entered. Then, when the TVA of teacher is calculated for teacher j, it will be a weighted mean of the teacher's true TVA and the TVAs of teachers of any students with mis-attributed IDs. Therefore,

$$E(\widehat{TVA_j}) = \frac{p}{e(1-p)+p(1-e)+ep} TVA_j + \frac{e}{e(1-p)+p(1-e)+ep} \overline{TVA_j},$$

where $\overline{TVA_j}$ is the mean TVA in the teacher population and $\widehat{TVA_j}$ is the estimate of the TVA for teacher j. This expression formalizes the intuition that the bias decreases in the true probability of switching p and increases in the error rate e.

Appendix D: Teacher Hiring Policy

The policy simulations in this paper resemble those of Staiger and Rockoff (2010). We simulate policies that fire the observed worst X percent of teachers (allowing X to be 5, 10, 50, 60, 70, and 80) after T years of observation (allowing T=1, 2, and 5).

We start by simulating TVA for a set of 100,000 teachers. Teachers' value-added in mean test scores is drawn from a distribution with mean 0 and standard deviation 0.148 (consistent with our estimates in Table 2). The initial population of teachers are assigned integer experience levels, which are drawn with equal probabilities from 0-30 years. We decrease a teacher's value-added by 0.08 if she has 0 or 1 years of experience (consistent with our estimates of the inexperience penalty in Table 3). We allow teachers to work for 30 years before they retire and are replaced with new teachers drawn from the normal distribution with mean 0 and variance 0.148. We then simulate the policy, allowing the distribution of employed teachers' quality and experience levels to change over time.

Each year, we assign each teacher 28 new students (the average public school class size in our data). We simulate each student's mean test scores as a linear function of mean TVA, a time-variant classroom specific shock, and a time-variant, idiosyncratic, individual shock. We ignore school-level, time invariant shocks since these are perfectly observable to policy makers with enough years of data. Following our calculations in Section 4, we assume these shocks are normally distributed, with mean 0 and the variances documented in Table 2. Consistent with a probation period of T years, after T years of observation, we estimate a teacher's observed TVA based on their students' test scores. Teachers whose observed TVA is below the X percentile of the observed TVA distribution are not confirmed and are replaced with new, inexperienced teachers with TVA drawn from $\mathcal{N}(0, 0.148)$.

To solve for the steady state effect of this policy on mean test scores, we simulate the policy in each year until $\epsilon = |\overline{TVA_t} - \overline{TVA_{t-1}}|/|\overline{TVA_{t-1}}| < 0.0005$. Additionally, to provide insight into how long it would take the policy to achieve its steady state effects, we report the mean TVA's of teachers under different policies at t=2, 5, and 15 years. Since the means in a given year may be affected by particularly bad or good draws of new teachers, we bootstrap these values by running the policy simulation 200 times for each policy and report the 95% confidence interval for the mean TVA values. We typically find that these confidence intervals are quite tight.

Appendix Tables

Table A1: Teacher-Level Summary Statistics

	(1)	(2)	(3)	(4)	(5)	(6)
	Government			Private		
	Mean	SD	N	Mean	SD	N
Female	0.449	0.497	3,829	0.768	0.422	4,733
Local	0.273	0.445	3,827	0.538	0.499	4,731
Some Training	0.904	0.294	3,829	0.221	0.415	4,731
BA Plus	0.514	0.500	3,829	0.255	0.436	4,734
Mean Salary	7,671 ($129)	3,746 ($63)	3,829	1,407 ($24)	997 ($17)	4,731
Multigrade	0.130	0.287	1,756	0.115	0.298	1,346
Temporary Contract	0.229	0.420	3,824	0.838	0.368	4,646
Year Started	1,990.80	10.710	3,432	2,002.17	7.749	3,159
Mean Days Absent	2.644	3.297	3,825	1.936	3.368	4,728
Mean Teacher Test Score	3.041	0.569	1,175	2.861	0.606	1,046
Mean School Basic Facilities	-0.473	0.743	3,667	0.562	1.114	4,651
Mean School Extra Facilities	-0.607	1.401	3,686	0.716	1.033	4,697
Mean Student Household Assets	-0.236	0.812	1,699	0.484	1.022	1,311
Mean Student Mother Primary Education	0.298	0.242	1,699	0.467	0.287	1,311
Mean Student Father Primary Education	0.580	0.245	1,699	0.739	0.242	1,311
Mean Change in Math Scores	0.393	0.499	1,533	0.355	0.488	975
Year 2 - Year 1	0.206	0.647	557	0.226	0.546	322
Year 3- Year 2	0.438	0.463	662	0.511	0.403	316
Years 4 - Year 3	0.475	0.561	1,041	0.354	0.490	573
Mean Change in English Scores	0.393	0.475	1,533	0.338	0.461	975
Year 2 - Year 1	0.303	0.652	557	0.187	0.459	322
Year 3- Year 2	0.375	0.454	662	0.408	0.402	316
Years 4 - Year 3	0.462	0.530	1,041	0.389	0.490	573
Mean Change in Urdu Scores	0.444	0.453	1,533	0.423	0.434	975
Year 2 - Year 1	0.306	0.633	557	0.317	0.459	322
Year 3- Year 2	0.444	0.424	662	0.497	0.368	316
Years 4 - Year 3	0.533	0.502	1,041	0.445	0.451	573
Mean Change in Mean Scores	0.410	0.413	1,533	0.372	0.399	975
Year 2 - Year 1	0.272	0.575	557	0.243	0.411	322
Year 3- Year 2	0.419	0.372	662	0.472	0.327	316
Years 4 - Year 3	0.490	0.461	1,041	0.396	0.409	573

This table presents teacher-level summary statistics across 4 rounds of the LEAPS survey (2004-2007). Changes in test scores are calculated by averaging over the difference between a student's test scores in time t and time $t-1$. Household assets and school basic and extra facilities are predicted from a principal components analysis of indicator variables for the presence of different assets, and school facilities and are normalized by year observed. The household asset measure is the first factor of a principal components analysis of indicator variables for ownership of beds, a radio, a television, a refrigerator, a bicycle, a plow, agricultural tools, tables, fans, a tractor, cattle, goats, chicken, watches, a motor rickshaw, a scooter, a car, a telephone, and a tubewell following methods discussed by Filmer and Pritchett (2001). The two indices for school facilities are constructed as the first predicted component from principles components analyses of indicator variables for "basic" and "extra" school facilities. Extra school facilities consist of a library, computer, sports, hall, school wall, fans, and electricity. The basic facilities consist of whether the school has desks/chairs as a seating arrangement, blackboards per child, toilets per child and classrooms per child.

Table A2: School-Level Summary Statistics

	(1)	(2)	(3)	(4)	(5)	(6)
	Government			Private		
	Mean	SD	N	Mean	SD	N
Number of Students	197.081	194.749	1,956	153.101	111.968	1,185
Number of Teachers	5.669	5.055	1,956	7.311	4.055	1,185
Percent Teachers with a BA	0.404	0.296	1,956	0.236	0.200	1,184
Percent Teachers with Some Training	0.915	0.174	1,956	0.280	0.236	1,176
Student/Teacher Ratio	35.599	14.445	1,913	21.171	8.963	1,159
Library	0.223	0.417	1,956	0.380	0.486	1,185
Computer	0.010	0.098	1,956	0.262	0.440	1,185
Sports	0.109	0.312	1,956	0.345	0.476	1,185
Hall	0.070	0.254	1,956	0.186	0.390	1,185
Wall	0.656	0.475	(student test score) 1,956	0.956	0.205	1,185
Fans	0.474	0.499	1,952	0.931	0.254	1,181
Electricity	0.539	0.499	1,956	0.948	0.223	1,184

This table presents school-level summary statistics across 4 rounds of the LEAPS survey (2004-2007). Each observation is at the school-year level.

Table A3: Variation in Grades Taught by Teachers in the Public Sector and the Number of Times Teachers are Observed

	(1) Observed Once	(2) Observed Twice	(3) Observed Three Times	(4) Observed Four Times
Only Grade 3	235	37	14	14
Restricted Sample	235	33	8	11
Only Grade 4	166	14	1	0
Restricted Sample	166	12	0	0
Only Grade 5	148	15	0	0
Restricted Sample	148	8	0	0
Grades 3 and 4	31	235	53	12
Restricted Sample	31	31	1	0
Grades 3 and 5	13	37	8	0
Restricted Sample	13	35	5	0
Grades 4 and 5	25	110	18	0
Restricted Sample	25	26	1	0
Grades 3, 4, 5	8	48	214	83
Restricted Sample	8	14	9	1

This table reports counts of the number of teachers who are observed teaching only Grade 3, only Grade 4, only Grade 5, only Grades 3 and 4, only Grades 4 and 5, and Grades 3, 4, and 5 by how many times they were observed in the public sector. The restricted sample excludes teachers who are ever observed teaching two classes of students who appear to be the same in two consecutive years (25 percent or more of the students in year t were taught by the same teacher in year $t-1$).

Table A4: Correlations Between Teacher Test Scores and Teacher Characteristics in the Sample of Public School Teachers with Test Scores

	(1) Math	(2) Math	(3) English	(4) English	(5) Urdu	(6) Urdu	(7) Mean	(8) Mean
Female	-0.252***	N.A.	-0.110*	N.A.	-0.116***	N.A.	-0.159***	N.A.
	(0.042)		(0.057)		(0.039)		(0.037)	
Local	0.021	-0.047	0.006	-0.089	-0.006	-0.026	0.007	-0.054
	(0.043)	(0.090)	(0.063)	(0.115)	(0.043)	(0.085)	(0.038)	(0.070)
Some Teacher Training	0.311	0.222	0.281	0.226	0.107	-0.037	0.233	0.137
	(0.211)	(0.300)	(0.214)	(0.288)	(0.134)	(0.180)	(0.150)	(0.197)
Has BA or Better	0.246***	0.225***	0.312***	0.268***	0.201***	0.154**	0.253***	0.216***
	(0.051)	(0.085)	(0.061)	(0.094)	(0.042)	(0.072)	(0.039)	(0.061)
Had >3 *Years of Exp in* 2007	-0.040	0.023	0.092	-0.002	0.071	0.206	0.041	0.075
	(0.099)	(0.184)	(0.091)	(0.190)	(0.069)	(0.130)	(0.072)	(0.134)
Temporary Contract	-0.111*	0.110	0.212***	0.369***	0.013	0.185	0.038	0.221**
	(0.064)	(0.135)	(0.068)	(0.140)	(0.056)	(0.116)	(0.049)	(0.102)
Fixed Effects	District	School	District	School	District	School	District	School
Number of Observations	1,105	1,105	1,105	1,105	1,105	1,105	1,105	1,105
Adjusted R Squared	0.070	0.042	0.062	0.167	0.049	0.114	0.085	0.200
Within-Adjusted R Squared	0.070	0.033	0.064	0.076	0.037	0.029	0.085	0.080
F	19.125	2.637	15.886	5.406	8.121	2.508	20.595	5.979
Clusters	491	491	491	491	491	491	491	491

This table reports estimates of the association between teachers' test scores on the same test given to students and teacher characteristics. The association between female and content knowledge in the public sector cannot be credibly estimated in the presence of school fixed effects because the public sector is not co-educational. Very few public schools (29) are observed with both male and female teachers over the course of the sample. Observations are at the teacher level and characteristics are time invariant. In cases where a teacher was tested more than once, the outcome variables are the average across multiple test scores. All regressions include district (odd columns) or school (even columns) fixed effects. The Within-Adjusted R Squared reports the adjusted R Squared within districts (odd columns) or schools (even columns). Standard errors are clustered at the school level. *, **, and *** indicate significance at the 10, 5, and 1 percent levels.

Table A5: Relationship Between Mean TVA and Teacher Characteristics for the Sample of Tested Teachers

	(1) Mean TVA	(2) Mean TVA	(3) Mean TVA	(4) Mean TVA
Female	0.064**	N.A.	0.048*	N.A.
	(0.026)		(0.028)	
Local	0.024	-0.001	0.031	-0.057
	(0.028)	(0.049)	(0.033)	(0.106)
Some Teacher Training	-0.096	-0.207*	-0.107	-0.489
	(0.074)	(0.123)	(0.097)	(0.386)
Has BA or Better	0.028	0.009	0.006	-0.050
	(0.032)	(0.057)	(0.037)	(0.121)
Had >3 *Years of Exp in* 2007	0.041	0.148	-0.021	0.150
	(0.047)	(0.094)	(0.059)	(0.336)
Temporary Contract	-0.013	0.064	0.015	0.077
	(0.043)	(0.081)	(0.050)	(0.176)
Fixed Effects	District	School	District	School
Number of observations	919	919	622	622
Adjusted R Squared	0.217	0.418	0.229	0.291
F	1.523	0.793	1.103	0.661
Clusters	469	469	439	439

This table reports estimates of the association between TVA and teacher characteristics for the samples of teacher for whom test scores are available. Columns 1 and 2 include the set of teachers for whom there are any test scores, and Columns 3 and 4 include those who were tested in at least two different years. The association between female and TVA in the public sector cannot be credibly estimated in the presence of school fixed effects because the public sector is not co-educational. Very few public schools (29) are observed with both male and female teachers over the course of the sample. The F-statistic is for a F-test of all the covariates. All regressions include district (odd columns) or school (even columns) fixed effects. Standard errors are clustered at the school level. *, **, and *** indicate significance at the 10, 5, and 1 percent levels.

Table A6: Correlation Between TVA Specifications in the Public Sector

	(1)	(2)	(3)	(4)
	SES and School Input Controls		**+Classroom Average Controls**	
	Across Schools	Within Schools	Across Schools	Within Schools
English	0.946	0.907	0.768	0.749
Math	0.913	0.898	0.774	0.788
Urdu	0.911	0.885	0.703	0.685

Columns 1 and 2 report the correlations between the baseline TVA estimates in English, math, and Urdu, and TVA estimates that are calculated controlling for gender, age, household assets, basic and extra school facilities indices, mother and father education, whether the classroom is multigrade, and student-teacher ratios. The household asset measure is the first factor of a principal components analysis of indicator variables for ownership of beds, a radio, a television, a refrigerator, a bicycle, a plow, agricultural tools, tables, fans, a tractor, cattle, goats, chicken, watches, a motor rickshaw, a scooter, a car, a telephone, and a tubewell following methods discussed by Filmer and Pritchett (2001). The two indices for school facilities are constructed as the first predicted component from principles components analyses of indicator variables for "basic" and "extra" school facilities. Extra school facilities consist of a library, computer, sports, hall, school wall, fans, and electricity. The basic facilities consist of whether the school has desks/chairs as its seating arrangement, blackboards per child, toilets per child and classrooms per child. Columns 3 and 4 report the correlations between the baseline TVA estimates and estimates that additionally include controls for classroom-level mean lagged test scores and mean household assets. Odd columns report the correlations for the across-school TVA estimates, while even columns report the correlations for the within-school TVA estimates.

Table A7: Variation in Grades Taught by Teachers in the Private Sector and the Number of Times Teachers are Observed

	(1) Observed Once	(2) Observed Twice	(3) Observed Three Times	(4) Observed Four Times
Only Grade 3	347	35	2	0
Restricted Sample	347	32	1	0
Only Grade 4	275	8	0	0
Restricted Sample	275	8	0	0
Only Grade 5	166	27	0	0
Restricted Sample	166	10	0	0
Grades 3 and 4	29	83	19	6
Restricted Sample	29	11	1	0
Grades 3 and 5	11	15	6	0
Restricted Sample	11	10	1	0
Grades 4 and 5	26	31	19	0
Restricted Sample	26	10	0	0
Grades 3, 4, 5	3	28	25	28
Restricted Sample	3	6	2	0

This table reports counts of the number of private school teachers who are observed teaching only Grade 3, only Grade 4, only Grade 5, only Grades 3 and 4, only Grades 4 and 5, and Grades 3, 4, and 5 by how many times they were observed. The restricted sample excludes teachers who are ever observed teaching two classes of students who appear to be the same in two consecutive years (25 percent or more of the students in year t were taught by the same teacher in year $t-1$).

Table A8: Sources of Variation in Teacher Value-Added Calculations for the Private Sector

	(1) Number of Teachers	(2) Number of Students	(3) Teachers in Schools With >1 Teacher With Tested Students	(4) Students in Schools With >1 Teachers With Tested Students
Round 1	303	3,617	0	0
Round 2	336	3,340	97	846
Round 3	579	6,777	524	6,247
Round 4	599	5,911	478	5,020

This table presents the breakdown of the data used to calculate within- and across-school TVAs for the private sector. Within-school TVAs require teachers to teach in schools where more than one teacher has tested students (so that the mean school effect is not equal to the sole teacher's TVA). The sample of students driving variation in the within-school TVAs are the students who attend schools where more than one teacher has tested students.

Table A9: Effect of a 1 SD Improvement in School, Teacher, Classroom, and Individual Effects in the Private Sector

	(1) **Math**	(2) **Urdu**	(3) **English**	(4) **Average**
		Panel A: Full Sample		
Classroom	0.318	0.281	0.232	0.277
School	0.099	0	0.142	0.080
Teacher	0.067	0.128	0.094	0.096
Individual	0.442	0.426	0.393	0.420
		Panel B: Restricted Sample		
Classroom	0.340	0.224	0.226	0.263
School	0	0	0.122	0.041
Teacher	0.125	0.214	0.141	0.160
Individual	0.440	0.419	0.399	0.419

This table reports the effect of receiving a 1sd higher classroom, school, teacher, or individual idiosyncratic shock on students' subject-level test scores, as well as the average effect across the three, in the private sector. Test scores are estimated with IRT. To arrive at these numbers, we use equation (1) to estimate teacher-year fixed effects in the panel dataset of student test scores. Denote $\widehat{\delta_{jst}}$ as the teacher-year fixed effect for teacher j in school s in year t. Then, the school variance is $Cov(\widehat{\delta_{jst}}, \widehat{\delta_{j'st}})$, the teacher variance is $Cov(\widehat{\delta_{jst}}, \widehat{\delta_{jst'}}) - Cov(\widehat{\delta_{jst}}, \widehat{\delta_{j'st}})$, the classroom variance is the variance of $\widehat{\delta_{j'st}}$ minus the sampling bias, which we solve for analytically in Appendix B, and the individual variance is the variance of the residuals. In Panel A, the sample includes all students and teachers in public schools. In Panel B, the restricted sample excludes teachers who are ever observed teaching two classes of students who appear to be the same in two consecutive years (25 percent or more of the students in year t were taught by the same teacher in year $t-1$).

Table A10: Do Student Test Score Trends Predict Being Taught by a Contract Teacher?

	(1) Mean Test Scores	(2) Mean Test Scores	(3) Had a Contract Teacher
$Year$	0.134***	0.145***	
	(0.013)	(0.013)	
$I(Received\,Contract\,Teacher)$	0.048	0.069	
	(0.078)	(0.083)	
$Year \times I(Received\,Contract\,Teacher)$	-0.015	-0.011	
	(0.023)	(0.024)	
$Mean\,Test\,Score\,Gain$			-0.014
			(0.016)
District FE	Y	Y	Y
School FE	N	Y	N
Grade by Lagged Test Score Interactions	Y	Y	N
Number of Observations	25,296	25,296	15,956
Adjusted R Squared	0.637	0.677	0.037
Clusters	478	478	497

This table tests whether better students are allocated to contract teachers. The first column estimates trends in student test scores before the receipt of a contract teacher in schools that did and did not receive contract teachers by regressing mean test scores on a continuous measure of the survey year, an indicator variable equal to 1 if the school received a contract teacher in the future, and their interaction, controlling for district fixed effects. The sample is restricted to schools that had not yet received the contract teacher. The next column compares the test score trends of students within schools who did or did not receive contract teachers before the receipt of the contract teacher by regressing mean test scores on a continuous measure of the survey year, an indicator variable equal to 1 if the student received a contract teacher in the future, and their interaction, controlling for district fixed effects. The sample is restricted to students who had not yet been taught by a contract teacher. In Columns 1 and 2, an observation is a student-year. The final regression regresses an indicator for whether a student ever had a contract teacher on their mean test score gains (residualized by testing round and grade) in the years prior to receiving a contract teacher, controlling for district fixed effects. In this sample, each student is observed once. Standard errors are clustered at the school level. *, **, and *** indicate significance at the 10, 5, and 1 percent levels.

Table A11: The Effect of Teacher Contract Status on Mean TVA

	(1) Mean TVA	(2) SE	(3) Boot-Strapped P-value	(4) N	(5) Within School Mean TVA	(6) SE	(7) Boot-Strapped P-value	(8) N
RD (2 Year)	0.840	0.545	0.075	209	0.593	0.452	0.170	193
RD (3 Year)	0.293	0.262	0.246	345	0.353*	0.179	0.074	320
RD (4 Year)	0.261	0.201	0.186	359	0.288**	0.134	0.039	334
RD (5 Year)	-0.047	0.105	0.630	608	0.086	0.064	0.356	576
RD (6 Year)	-0.000	0.096	0.966	634	0.085	0.060	0.305	601
RD (7 Year)	-0.001	0.094	0.971	635	0.080	0.058	0.331	602
RD (8 Year)	0.039	0.091	0.626	687	0.096*	0.057	0.206	651
RD (9 Year)	-0.044	0.085	0.549	766	0.044	0.058	0.672	725
RD (10 Year)	-0.036	0.081	0.572	801	0.039	0.053	0.685	758

This table reports the fuzzy RD effects of temporary contract status on mean TVA for different bandwidths (2-10 years), as well as p-values from cluster bootstrapped standard errors, which account for estimation error in the TVA's. Contract status is instrumented for with an indicator variable for whether a teacher was hired after 1998. All regressions contain linear time trends in month hired, which are allowed to differ before and after the budgetary shock, and district fixed effects. Regressions also include controls for student teacher ratios, student socioeconomic status, and whether a classroom is multigrade. Cross-school TVA estimates are de-meaned at the district-level. Observations are at the teacher level, and standard errors are clustered at the month hired level in the regression discontinuity specifications. *, **, and *** indicate significance at the 10, 5, and 1 percent levels according to the analytic standard errors.

Appendix Figures

Figure A1: Basic Pay Scale for Pakistani Civil Servants

BASIC PAY SCALE (BPS) CHART- 2016 (FINAL)

SCALE	YEAR	INIT	INCR	MAX	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
BPS-1	2015	6210	195	12060	6210	6405	6600	6795	6990	7185	7380	7575	7770	7965	8160	8355	8550	8745	8940	9135	9330	9525	9720	9915	10110	10305	10500	10695	10890	11085	11280	11475	11670	11865	12060
	2016	**7640**	**240**	**14840**	**7640**	**7880**	**8120**	**8360**	**8600**	**8840**	**9080**	**9320**	**9560**	**9800**	**10040**	**10280**	**10520**	**10760**	**11000**	**11240**	**11480**	**11720**	**11960**	**12200**	**12440**	**12680**	**12920**	**13160**	**13400**	**13640**	**13880**	**14120**	**14360**	**14600**	**14840**
BPS-2	2015	6335	220	12935	6335	6555	6775	6995	7215	7435	7655	7875	8095	8315	8535	8755	8975	9195	9415	9635	9855	10075	10295	10515	10735	10955	11175	11395	11615	11835	12055	12275	12495	12715	12935
	2016	**7790**	**275**	**16040**	**7790**	**8065**	**8340**	**8615**	**8890**	**9165**	**9440**	**9715**	**9990**	**10265**	**10540**	**10815**	**11090**	**11365**	**11640**	**11915**	**12190**	**12465**	**12740**	**13015**	**13290**	**13565**	**13840**	**14115**	**14390**	**14665**	**14940**	**15215**	**15490**	**15765**	**16040**
BPS-3	2015	6535	260	14335	6535	6795	7055	7315	7575	7835	8095	8355	8615	8875	9135	9395	9655	9915	10175	10435	10695	10955	11215	11475	11735	11995	12255	12515	12775	13035	13295	13555	13815	14075	14335
	2016	**8040**	**325**	**17790**	**8040**	**8365**	**8690**	**9015**	**9340**	**9665**	**9990**	**10315**	**10640**	**10965**	**11290**	**11615**	**11940**	**12265**	**12590**	**12915**	**13240**	**13565**	**13890**	**14215**	**14540**	**14865**	**15190**	**15515**	**15840**	**16165**	**16490**	**16815**	**17140**	**17465**	**17790**
BPS-4	2015	6730	300	15730	6730	7030	7330	7630	7930	8230	8530	8830	9130	9430	9730	10030	10330	10630	10930	11230	11530	11830	12130	12430	12730	13030	13330	13630	13930	14230	14530	14830	15130	15430	15730
	2016	**8280**	**370**	**19380**	**8280**	**8650**	**9020**	**9390**	**9760**	**10130**	**10500**	**10870**	**11240**	**11610**	**11980**	**12350**	**12720**	**13090**	**13460**	**13830**	**14200**	**14570**	**14940**	**15310**	**15680**	**16050**	**16420**	**16790**	**17160**	**17530**	**17900**	**18270**	**18640**	**19010**	**19380**
BPS-5	2015	6985	340	17185	6985	7325	7665	8005	8345	8685	9025	9365	9705	10045	10385	10725	11065	11405	11745	12085	12425	12765	13105	13445	13785	14125	14465	14805	15145	15485	15825	16165	16505	16845	17185
	2016	**8590**	**420**	**21190**	**8590**	**9010**	**9430**	**9850**	**10270**	**10690**	**11110**	**11530**	**11950**	**12370**	**12790**	**13210**	**13630**	**14050**	**14470**	**14890**	**15310**	**15730**	**16150**	**16570**	**16990**	**17410**	**17830**	**18250**	**18670**	**19090**	**19510**	**19930**	**20350**	**20770**	**21190**
BPS-6	2015	7235	375	18485	7235	7610	7985	8360	8735	9110	9485	9860	10235	10610	10985	11360	11735	12110	12485	12860	13235	13610	13985	14360	14735	15110	15485	15860	16235	16610	16985	17360	17735	18110	18485
	2016	**8900**	**470**	**23000**	**8900**	**9370**	**9840**	**10310**	**10780**	**11250**	**11720**	**12190**	**12660**	**13130**	**13600**	**14070**	**14540**	**15010**	**15480**	**15950**	**16420**	**16890**	**17360**	**17830**	**18300**	**18770**	**19240**	**19710**	**20180**	**20650**	**21120**	**21590**	**22060**	**22530**	**23000**
BPS-7	2015	7490	415	19940	7490	7905	8320	8735	9150	9565	9980	10395	10810	11225	11640	12055	12470	12885	13300	13715	14130	14545	14960	15375	15790	16205	16620	17035	17450	17865	18280	18695	19110	19525	19940
	2016	**9220**	**510**	**24520**	**9220**	**9730**	**10240**	**10750**	**11260**	**11770**	**12280**	**12790**	**13300**	**13810**	**14320**	**14830**	**15340**	**15850**	**16360**	**16870**	**17380**	**17890**	**18400**	**18910**	**19420**	**19930**	**20440**	**20950**	**21460**	**21970**	**22480**	**22990**	**23500**	**24010**	**24520**
BPS-8	2015	7750	455	21400	7750	8205	8660	9115	9570	10025	10480	10935	11390	11845	12300	12755	13210	13665	14120	14575	15030	15485	15940	16395	16850	17305	17760	18215	18670	19125	19580	20035	20490	20945	21400
	2016	**9540**	**560**	**26340**	**9540**	**10100**	**10660**	**11220**	**11780**	**12340**	**12900**	**13460**	**14020**	**14580**	**15140**	**15700**	**16260**	**16820**	**17380**	**17940**	**18500**	**19060**	**19620**	**20180**	**20740**	**21300**	**21860**	**22420**	**22980**	**23540**	**24100**	**24660**	**25220**	**25780**	**26340**
BPS-9	2015	8015	495	22865	8015	8510	9005	9500	9995	10490	10985	11480	11975	12470	12965	13460	13955	14450	14945	15440	15935	16430	16925	17420	17915	18410	18905	19400	19895	20390	20885	21380	21875	22370	22865
	2016	**9860**	**610**	**28160**	**9860**	**10470**	**11080**	**11690**	**12300**	**12910**	**13520**	**14130**	**14740**	**15350**	**15960**	**16570**	**17180**	**17790**	**18400**	**19010**	**19620**	**20230**	**20840**	**21450**	**22060**	**22670**	**23280**	**23890**	**24500**	**25110**	**25720**	**26330**	**26940**	**27550**	**28160**
BPS-10	2015	8275	544	24595	8275	8819	9363	9907	10451	10995	11539	12083	12627	13171	13715	14259	14803	15347	15891	16435	16979	17523	18067	18611	19155	19699	20243	20787	21331	21875	22419	22963	23507	24051	24595
	2016	**10180**	**670**	**30280**	**10180**	**10850**	**11520**	**12190**	**12860**	**13530**	**14200**	**14870**	**15540**	**16210**	**16880**	**17550**	**18220**	**18890**	**19560**	**20230**	**20900**	**21570**	**22240**	**22910**	**23580**	**24250**	**24920**	**25590**	**26260**	**26930**	**27600**	**28270**	**28940**	**29610**	**30280**
BPS-11	2015	8540	595	26390	8540	9135	9730	10325	10920	11515	12110	12705	13300	13895	14490	15085	15680	16275	16870	17465	18060	18655	19250	19845	20440	21035	21630	22225	22820	23415	24010	24605	25200	25795	26390
	2016	**10510**	**740**	**32710**	**10510**	**11250**	**11990**	**12730**	**13470**	**14210**	**14950**	**15690**	**16430**	**17170**	**17910**	**18650**	**19390**	**20130**	**20870**	**21610**	**22350**	**23090**	**23830**	**24570**	**25310**	**26050**	**26790**	**27530**	**28270**	**29010**	**29750**	**30490**	**31230**	**31970**	**32710**
BPS-12	2015	9055	650	28555	9055	9705	10355	11005	11655	12305	12955	13605	14255	14905	15555	16205	16855	17505	18155	18805	19455	20105	20755	21405	22055	22705	23355	24005	24655	25305	25955	26605	27255	27905	28555
	2016	**11140**	**800**	**35140**	**11140**	**11940**	**12740**	**13540**	**14340**	**15140**	**15940**	**16740**	**17540**	**18340**	**19140**	**19940**	**20740**	**21540**	**22340**	**23140**	**23940**	**24740**	**25540**	**26340**	**27140**	**27940**	**28740**	**29540**	**30340**	**31140**	**31940**	**32740**	**33540**	**34340**	**35140**
BPS-13	2015	9700	715	31150	9700	10415	11130	11845	12560	13275	13990	14705	15420	16135	16850	17565	18280	18995	19710	20425	21140	21855	22570	23285	24000	24715	25430	26145	26860	27575	28290	29005	29720	30435	31150
	2016	**11930**	**880**	**38330**	**11930**	**12810**	**13690**	**14570**	**15450**	**16330**	**17210**	**18090**	**18970**	**19850**	**20730**	**21610**	**22490**	**23370**	**24250**	**25130**	**26010**	**26890**	**27770**	**28650**	**29530**	**30410**	**31290**	**32170**	**33050**	**33930**	**34810**	**35690**	**36570**	**37450**	**38330**
BPS-14	2015	10340	790	34040	10340	11130	11920	12710	13500	14290	15080	15870	16660	17450	18240	19030	19820	20610	21400	22190	22980	23770	24560	25350	26140	26930	27720	28510	29300	30090	30880	31670	32460	33250	34040
	2016	**12720**	**980**	**42120**	**12720**	**13700**	**14680**	**15660**	**16640**	**17620**	**18600**	**19580**	**20560**	**21540**	**22520**	**23500**	**24480**	**25460**	**26440**	**27420**	**28400**	**29380**	**30360**	**31340**	**32320**	**33300**	**34280**	**35260**	**36240**	**37220**	**38200**	**39180**	**40160**	**41140**	**42120**
BPS-15	2015	10985	905	38135	10985	11890	12795	13700	14605	15510	16415	17320	18225	19130	20035	20940	21845	22750	23655	24560	25465	26370	27275	28180	29085	29990	30895	31800	32705	33610	34515	35420	36325	37230	38135
	2016	**13510**	**1120**	**47110**	**13510**	**14630**	**15750**	**16870**	**17990**	**19110**	**20230**	**21350**	**22470**	**23590**	**24710**	**25830**	**26950**	**28070**	**29190**	**30310**	**31430**	**32550**	**33670**	**34790**	**35910**	**37030**	**38150**	**39270**	**40390**	**41510**	**42630**	**43750**	**44870**	**45990**	**47110**
BPS-16	2015	12910	1035	43960	12910	13945	14980	16015	17050	18085	19120	20155	21190	22225	23260	24295	25330	26365	27400	28435	29470	30505	31540	32575	33610	34645	35680	36715	37750	38785	39820	40855	41890	42925	43960
	2016	**15880**	**1280**	**54280**	**15880**	**17160**	**18440**	**19720**	**21000**	**22280**	**23560**	**24840**	**26120**	**27400**	**28680**	**29960**	**31240**	**32520**	**33800**	**35080**	**36360**	**37640**	**38920**	**40200**	**41480**	**42760**	**44040**	**45320**	**46600**	**47880**	**49160**	**50440**	**51720**	**53000**	**54280**
BPS-17	2015	20680	1555	51780	20680	22235	23790	25345	26900	28455	30010	31565	33120	34675	36230	37785	39340	40895	42450	44005	45560	47115	48670	50225	51780										
	2016	**25440**	**1930**	**64040**	**25440**	**27370**	**29300**	**31230**	**33160**	**35090**	**37020**	**38950**	**40880**	**42810**	**44740**	**46670**	**48600**	**50530**	**52460**	**54390**	**56320**	**58250**	**60180**	**62110**	**64040**										
BPS-18	2015	25940	1950	64940	25940	27890	29840	31790	33740	35690	37640	39590	41540	43490	45440	47390	49340	51290	53240	55190	57140	59090	61040	62990	64940										
	2016	**31890**	**2400**	**79890**	**31890**	**34290**	**36690**	**39090**	**41490**	**43890**	**46290**	**48690**	**51090**	**53490**	**55890**	**58290**	**60690**	**63090**	**65490**	**67890**	**70290**	**72690**	**75090**	**77490**	**79890**										
BPS-19	2015	40155	2075	81655	40155	42230	44305	46380	48455	50530	52605	54680	56755	58830	60905	62980	65055	67130	69205	71280	73355	75430	77505	79580	81655										
	2016	**49370**	**2560**	**100570**	**49370**	**51930**	**54490**	**57050**	**59610**	**62170**	**64730**	**67290**	**69850**	**72410**	**74970**	**77530**	**80090**	**82650**	**85210**	**87770**	**90330**	**92890**	**95450**	**98010**	**100570**										
BPS-20	2015	46705	3050	89405	46705	49755	52805	55855	58905	61955	65005	68055	71105	74155	77205	80255	83305	86355	89405																
	2016	**57410**	**3750**	**109910**	**57410**	**61160**	**64910**	**68660**	**72410**	**76160**	**79910**	**83660**	**87410**	**91160**	**94910**	**98660**	**102410**	**106160**	**109910**																
BPS-21	2015	51885	3375	99135	51885	55260	58635	62010	65385	68760	72135	75510	78885	82260	85635	89010	92385	95760	99135																
	2016	**63780**	**4150**	**121880**	**63780**	**67930**	**72080**	**76230**	**80380**	**84530**	**88680**	**92830**	**96980**	**101130**	**105280**	**109430**	**113580**	**117730**	**121880**																
BPS-22	2015	55755	3960	111195	55755	59715	63675	67635	71595	75555	79515	83475	87435	91395	95355	99315	103275	107235	111195																
	2016	**68540**	**4870**	**136720**	**68540**	**73410**	**78280**	**83150**	**88020**	**92890**	**97760**	**102630**	**107500**	**112370**	**117240**	**122110**	**126980**	**131850**	**136720**																

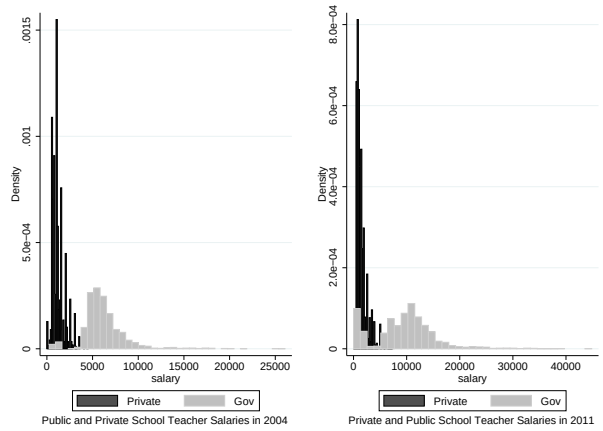
Muhammad Farooq
www.itechsoul.com

The columns of the figure give the initial salary and pay increments for an additional year of experience, and the rows denote the grade of the civil servant. Primary school teachers are grade 9, middle school teachers are grade 14, secondary school teachers are grade 16, and head teachers are grade 17. The figure is taken from ITECHSOUL (accessed Nov. 30, 2018).

Figure A2: Percent of Public School Teachers with a Bachelor's Degrees by Year Hired

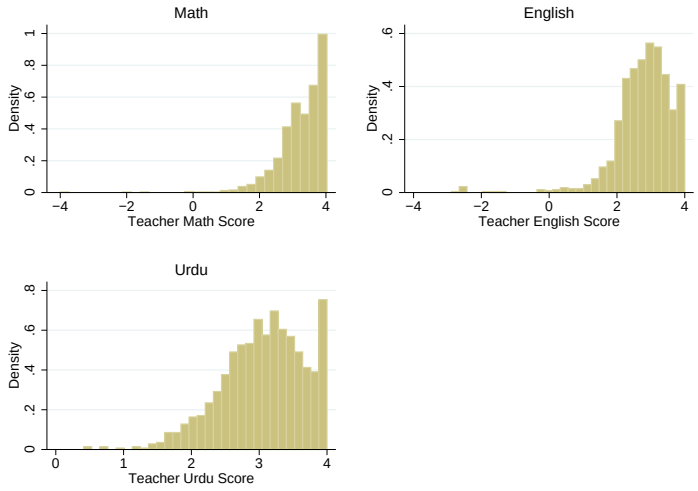
This figure plots the percent of teachers hired in a given year who had at least a bachelor's degree in the LEAPS sample.

Figure A3: Teacher Salaries in Public and Private Schools



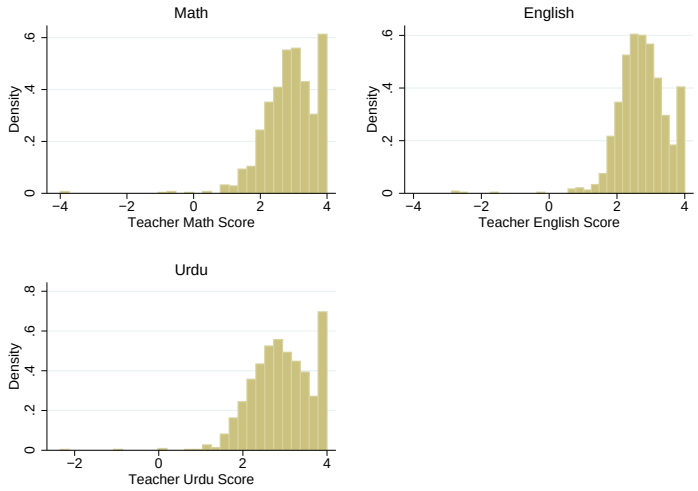
This figure plots the distribution of teacher salaries (in Rupees) in the private and public sector in our first year of data collection (2004) and in 2011.

Figure A4: Teacher Test Scores in Public Schools



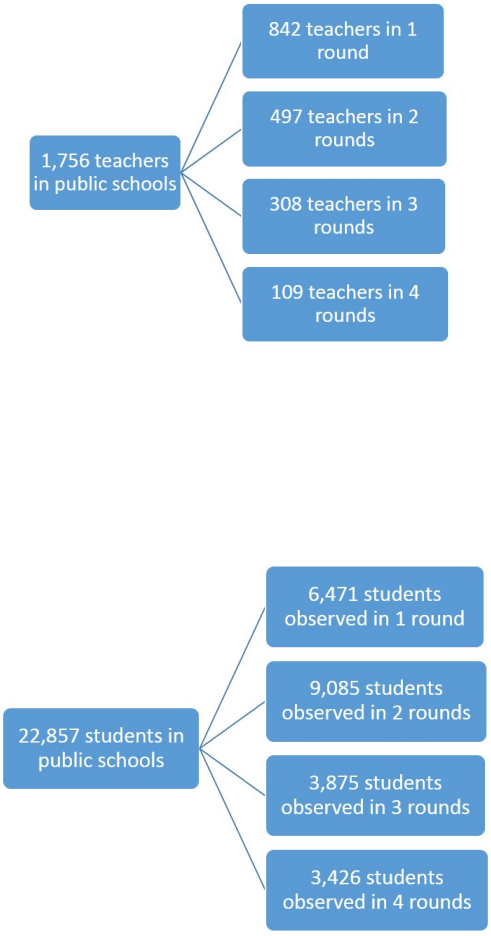
This figure plots the distribution of public school teachers' test scores on the same tests administered to students in math, English, and Urdu. Each teacher-year is treated as a separate observation.

Figure A5: Teacher Test Scores in Private Schools



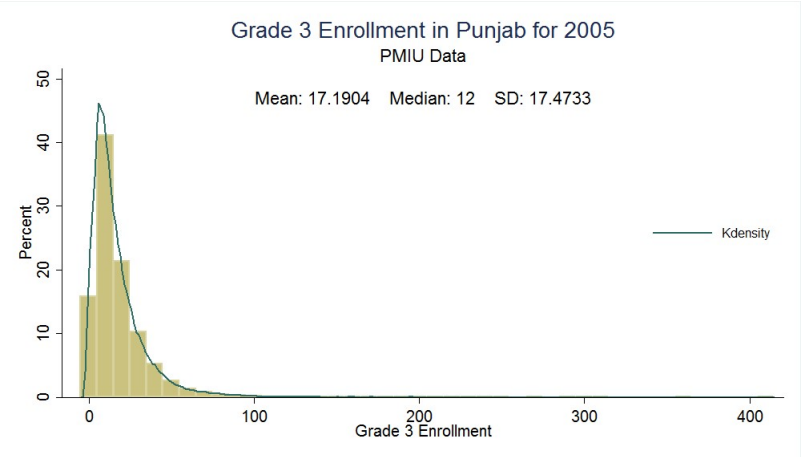
This figure plots the distribution of private school teachers' test scores on the same tests administered to students in math, English, and Urdu. Each teacher-year is treated as a separate observation.

Figure A6: Number of Rounds Public School Teachers and Students are Observed



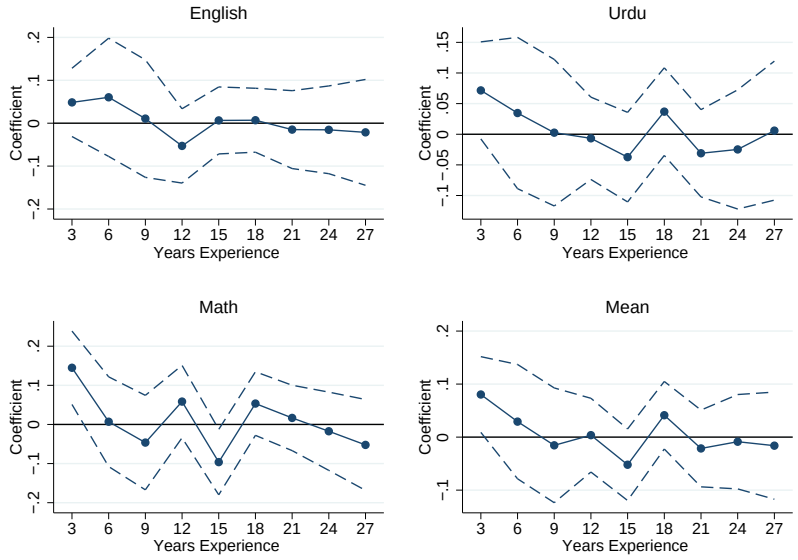
This figure reports the number of public school teachers and students in the LEAPS sample and the breakdown of how many teachers and students are observed in 1, 2, 3 and 4 rounds of data collection.

Figure A7: Third Grade Sizes in Public Schools in Punjab



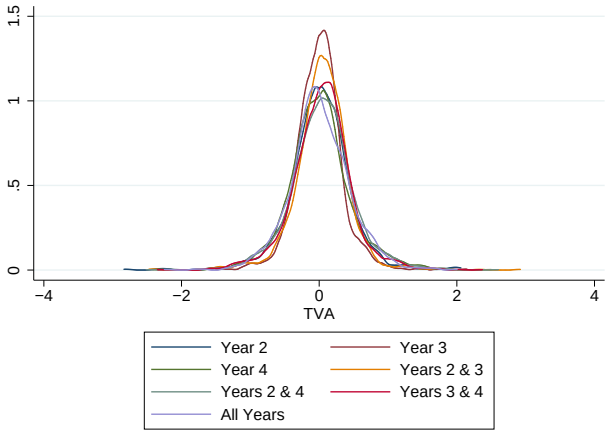
This figure reports the distribution of the number of students enrolled in third grade by school in Punjab using the Programme Monitoring and Implementation Unit data from 2005.

Figure A8: Non-Parametric Estimates of the Marginal Effect of Additional Years of Experience on Students' Learning in the Public Sector



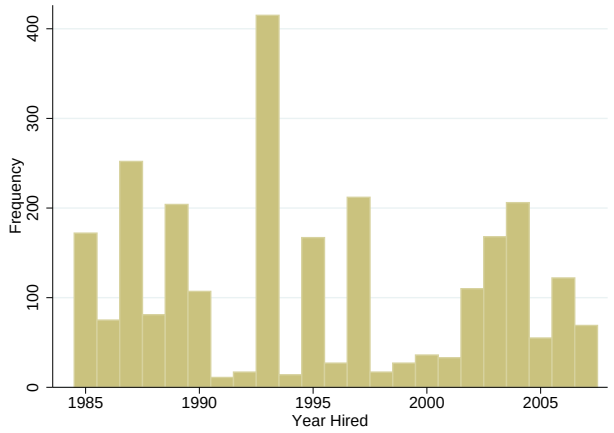
This figure reports estimates of the marginal treatment effects for each additional 3 years of teaching experience. The estimates are obtained by regressing student test scores on lagged student test scores interacted with child class fixed effects, year fixed effects, and indicator variables for having greater than or equal to 3 years of experience, 6 years of experience, and so on up to 27 years of experience. The sample was restricted to public school teachers with 30 or fewer years of experience.

Figure A9: Distributions of the Mean (Fixed Effect) TVA Estimates



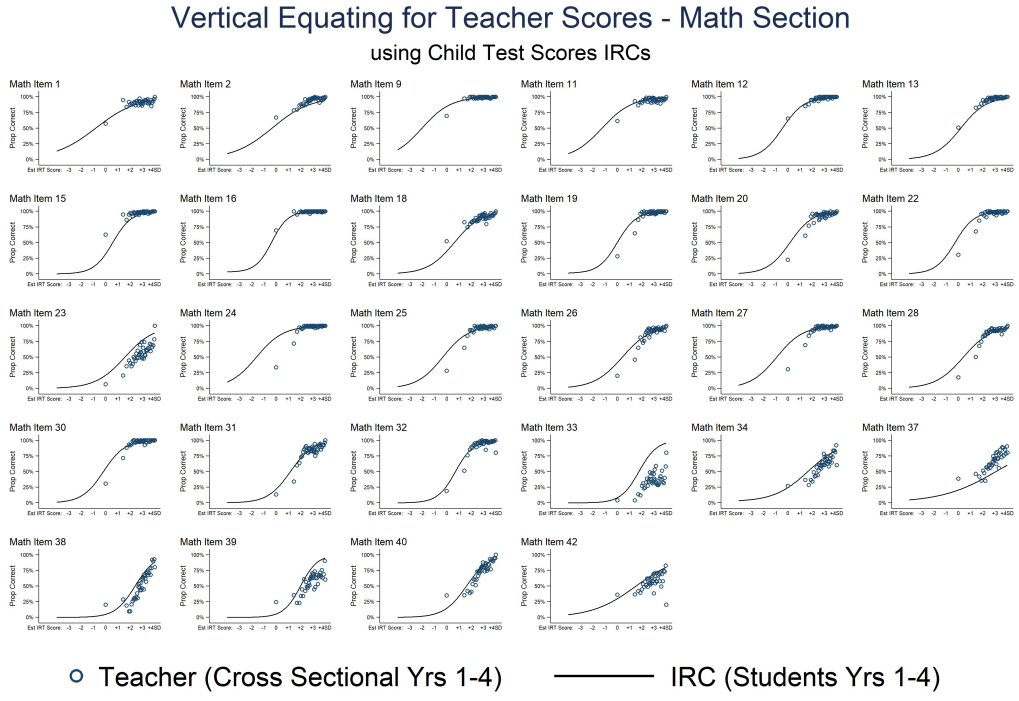
This figure plots the distribution of TVA estimates that were calculated using mutually exclusive random samples of data from year 2, year 3, year 4, years 2 and 3, years 3 and 4, years 2 and 4, and all three years. The samples are mutually exclusive to avoid mechanical correlations in the TVA estimates.

Figure A10: Number of Public Sector Teachers Hired by Year



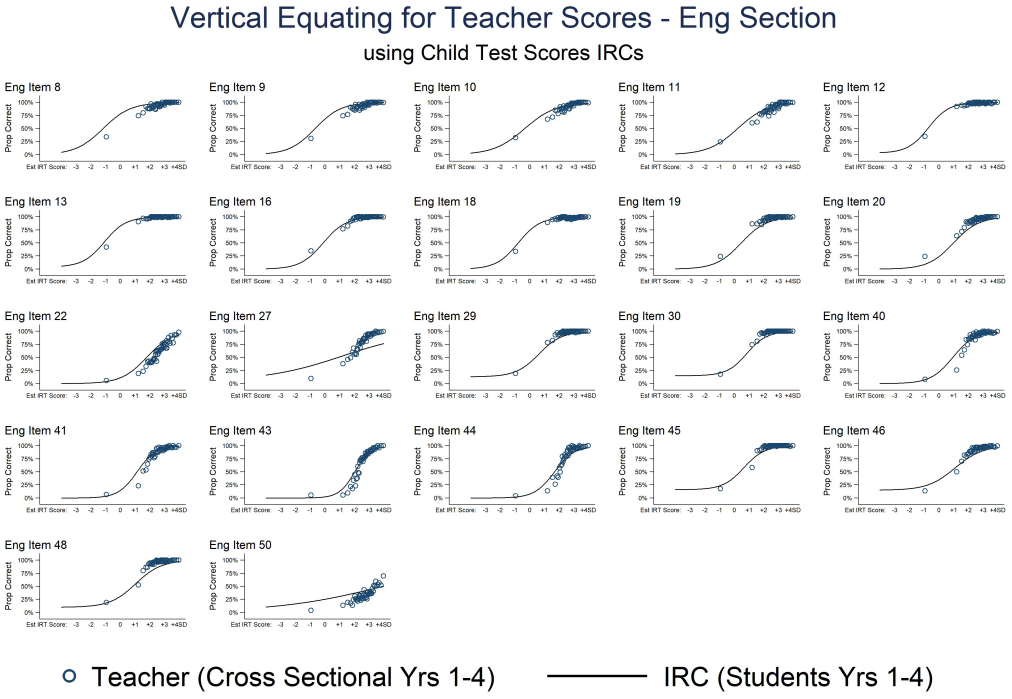
This figure plots the frequency of teachers hired by year in the public sector from 1985-2007.

Figure A11: Observed and Predicted Teacher Response by Test Question in Math



This figure plots the predicted portion of teachers who got each math item correct according to the IRT estimates (using item-level parameters estimated with the student data) against the actual portion of teachers who got that item correct in the data.

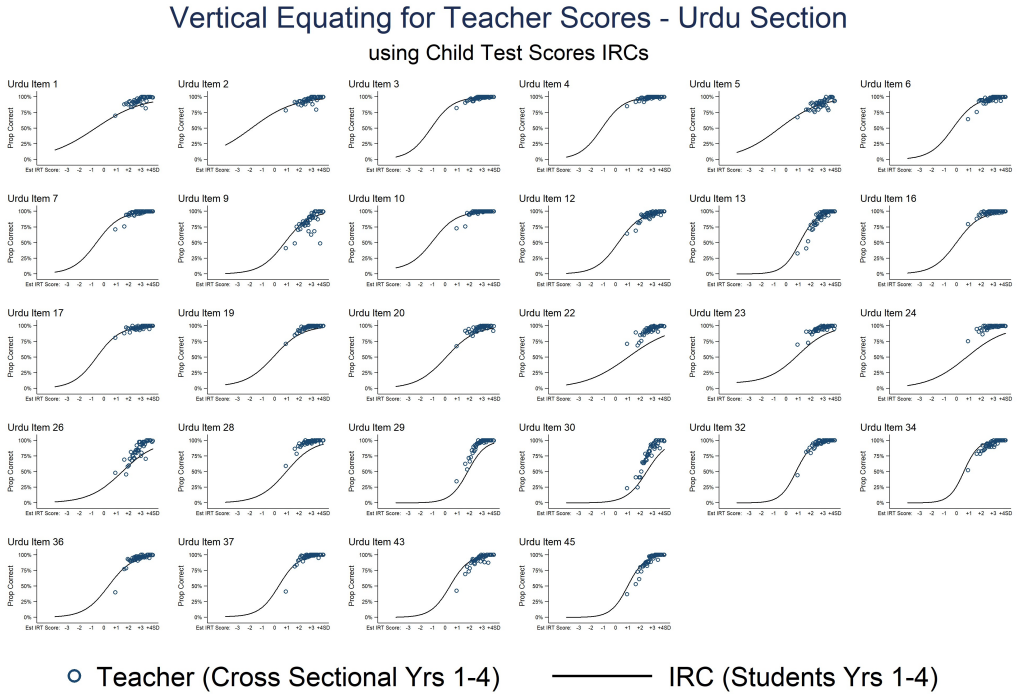
Figure A12: Observed and Predicted Teacher Response by Test Question in English



*Theta predicted using year 1-4 teacher data as cross-sectional

This figure plots the predicted portion of teachers who got each English item correct according to the IRT estimates (using item-level parameters estimated with the student data) against the actual portion of teachers who got that item correct in the data

Figure A13: Observed and Predicted Teacher Response by Test Question in Urdu



This figure plots the predicted portion of teachers who got each Urdu item correct according to the IRT estimates (using item-level parameters estimated with the student data) against the actual portion of teachers who got that item correct in the data

References

Das, Jishnu, and Tristan Zajonc. 2010. "India Shining and Bharat Drowning: Comparing Two Indian States to the Worldwide Distribution in Mathematics Achievement." *Journal of Development Economics*, 92(2): 175–187.

Filmer, Deon, and Lant Pritchett. 2001. "Estimating Wealth Effects Without Expenditure Data or Tears: An Application to Educational Enrollments in States of India." *Demography*, 38(1): 115–132.

ITECHSOUL. accessed Nov. 30, 2018. "Basic-Pay-Scales-BPS-Chart-2016-Final." *http://www.itechsoul.com/new-revised-pay-scale-chart-2016-for-government-employees-grade-01-to-22/basic-pay-scales-bps-chart-2016-final*.

Staiger, Douglas O, and Jonah E Rockoff. 2010. "Searching for Effective Teachers With Imperfect Information." *The Journal of Economic Perspectives*, 24(3): 97–117.