

The Misallocation of Pay and Productivity in the Public Sector: Evidence From the Labor Market for Teachers

Natalie Bau Jishnu Das

UToronto World Bank

June 14, 2017

How much should teachers be paid?

- How to recruit and reward teachers is key for educational policy
 - Teacher quality is consistently important for child learning and later life outcomes across studies.
 - Teacher salaries account for >80% of education budgets in most LMIC.
- Many policy makers argue that low salaries are *the* problem in recruitment.
 - Auguste et al. (2010): “Money was the most powerful lever in attracting and retaining top-third students.”
- Evidence from LMIC sparse because
 - Lack of teacher-student matched data.
 - Small schools: mean third grade size is 17 students in rural Pakistan.

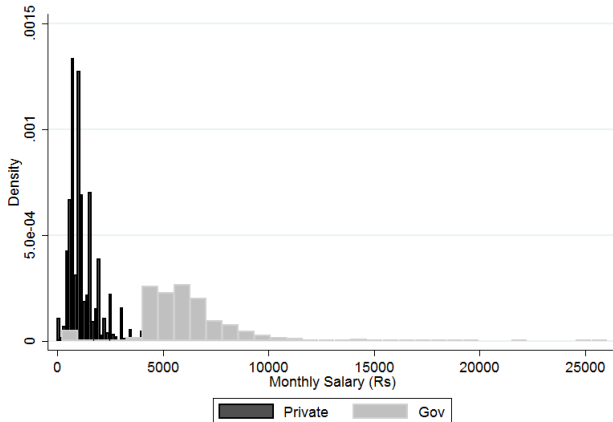
This Paper

- Use data we collected over 4 years in rural Pakistan to
 - Estimate the importance of teachers for child learning.
 - Estimate TVA for both public and private schools and argue for its validity in this context.
 - Examine correlates of TVA.
- Combine with policy experiment that suddenly shifted all hiring to lower-paid (35% lower) contract teachers to examine link between wages and TVA.
 - Slope: Correlation of wages and TVA in public and private schools.
 - Intercept: Did lower wages reduce TVA for immediate hires, as well as those hired after 4 years?

Relationship with the Literature

- TVA Estimation in the U.S. (Rockoff, 2004; Chetty et al., 2014; Kane and Staiger, 2008; Hanushek and Rivkin, 2012; Araujo et al., 2016)
 - Same methods with some variations and a novel test for validity that could have wide applicability in LMIC.
 - First set of results for private school teachers.
- Teachers wages in LMIC
 - Contract teacher experiments in LMIC with NGOs report higher learning at lower wages (Duflo et al., 2014; Muralidharan and Sundararaman, 2013).
 - Extend the validity of these findings to a large-scale policy change.
 - First RCT of doubling current teachers' wages in Indonesia find no effects on student learning (De Ree et al., 2015).
 - We study the extensive margin instead of the intensive margin.
- Public-private wage differences in OECD
 - Uses worker fixed-effects to identify wage premia ranging from 5-10%.
 - In LMIC contexts, this may be as high as 300-500%.

Teacher Salaries in 2004



Outline

- 1 **Data**
- 2 TVA Results
- 3 Regime Change Results
- 4 Conclude

LEAPS Data

Two key surveys in 112 villages of Punjab Province, Pakistan, each conducted every year from 2004-2007:

- Geo-coded survey of the universe of schools.
 - 574 public schools (1,533 teachers) and 345 private schools (975 teachers) in 112 villages.
 - Data on school and teacher characteristics.
 - Data on teacher test scores.
- Surveys of 3rd-6th grade children in the schools, including low-stakes test scores in math, Urdu, and English.
 - 22,857 children in public schools.
 - 9,741 children in private schools.

Test Scores

- Tests were administered by our team in each year in English, mathematics, and Urdu.
 - Supervised by team with clear instructions not to interfere.
 - Test booklets retrieved after class (no missing test material).
- Tests were “low-stakes”: teachers and students had no incentive to cheat.
- Use item response theory to grade tests.
 - “Weights” different items according to difficulty.
 - Item response theory allows us to equate tests over years using “linking items.”
- A subset of teachers were also tested.

[▶ Learning Over Time](#)[▶ Sample Questions](#)[▶ Teacher Test Scores](#)

Outline

- 1 Data
- 2 **TVA Results**
- 3 Regime Change Results
- 4 Conclude

TVA Empirical Strategy

$$\text{Child learning} = \text{past learning} + \text{classroom quality} + \text{year shock} \\ + \text{grade shock} + \text{idiosyncratic student shock}$$

- To back out *classroom quality*, estimate the effect of belonging to a classroom on test scores, controlling for
 - Past test scores.
 - Year fixed effects.
 - Grade fixed effects.
- *classroom quality* combines the effect of time invariant teacher quality and the effects of classroom components (e.g. peers).
- To get the effect of belonging to a 1 sd better classroom, estimate the variance of *classroom quality*.
- To get the effect of having a 1 sd better teacher, estimate the covariance between *classroom qualities* for the same teacher over time.

TVA Empirical Strategy

- Using the same framework, estimate *teacher quality* (TVA).
- Intuitively, portion of the classroom quality that does not change over time.

Variance of Teacher and Classroom Effects

- Classroom effect: the effect of being in a 1 sd better classroom on test scores.
- Teacher effect: the effect of having a 1 sd better teacher on test scores.
- Following Araujo et al. (2016), estimate classroom and teacher effects.
 - Some statistical innovations to account for the fact only 1 teacher is observed in a school-year.

	Math		English		Urdu		Average	
	Class	Teacher	Class	Teacher	Class	Teacher	Class	Teacher
Full Sample	0.321	0.258	0.300	0.190	0.312	0.184	0.311	0.211
Public Schools Only	0.356	0.199	0.337	0.134	0.351	0.152	0.348	0.162

Comparisons to the Literature

	Math	Vernacular
Punjab (Public Only)	0.20	0.15
Los Angeles (Staiger and Rockoff, 2010)	0.19	0.14
New York (Staiger and Rockoff, 2010)	0.15	0.12
Large Urban School District (Chetty et al., 2014)	0.13	0.10
Ecuador (Araujo et al., 2016)	0.09	0.09

- Higher end of still substantial variance in teacher quality in the U.S.

Association Between Observable Characteristics and Teacher Quality

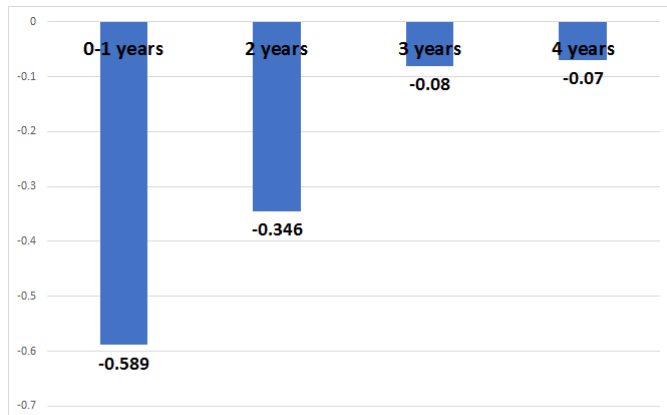
- First, estimate a single measure of teacher productivity (TVAs), equivalent to teacher “fixed effects.”
- Then, estimate the association between TVA and teacher characteristics.
- To estimate the effect of experience, use panel data to estimate the *within teacher* effect of a year of experience.

Association Between Teacher Characteristics and TVA

For both private and public school teachers,

- Education and teacher training are not associated with teacher productivity.
- Besides experience, only content knowledge has a strong association
 - Unadjusted for measurement error: 1 sd higher mean teacher scores \rightarrow 0.090*** (0.038) higher TVA.
 - Adjusted for measurement error: 1 sd higher mean teacher scores \rightarrow 0.298*** (0.072) higher TVA.
 - Consistent with findings of Bold et al. (2017) in Africa.

Effect of Years of Experience (Relative to 5+ Years)



Note: Associations for private school teachers are similar.

TVA Robustness

- Results rely on the assumption that lagged student test scores capture selection of students to teachers.
- We can test this:
 - When a student switches schools, does a student's future teacher's TVA predict his present teacher's TVA (Rothstein, 2010)?
 - No: Coefficient on mean TVA is 0.002 (0.046).
 - Are TVA estimates predictive of actual student gains when students switch schools (Chetty et al., 2014)?
 - Yes: Coefficient on mean TVA is 0.852*** (0.078).

What do we know?

	Pakistan	US	Ecuador*	Uganda**
Teachers are important	Y	Y	Y	Y
Teachers are heterogeneous	Y	Y	Y	Y
Characteristics explain little of TVA	Y	Y	Y	Y
Early experience matters	Y	Y	Y	N.A.
Content knowledge matters	Y	No Evidence	No Evidence	N.A.

*Araujo et al. (2016)

**Buhl-Wiggers et al. (2017)

Association Between TVA and Teacher Salaries



Note: TVA is measured in student test score sd.

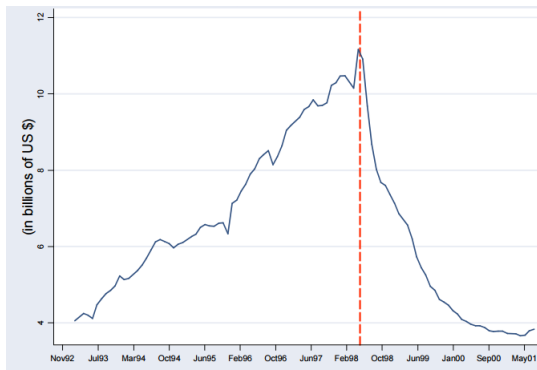
Outline

- ① Data
- ② TVA Robustness & Robustness
- ③ **Regime Change Results**
- ④ Conclude

Does Teacher Quality Respond to Salaries?

- Our TVA results suggest that there is little link between teacher salaries and teacher quality.
- Raises an important policy question: How would lowering teacher salaries affect the quality of teachers?
- A regime change following Pakistan's unexpected nuclear tests in 1998 allows us to look at the joint effect of a salary decrease combined with greater accountability.
- Compare teachers right hired right before and after the shock to capture the effect of being hired on a temporary contract.
- Complements work by De Ree et al. (2015) on teacher salaries in Indonesia.
 - De Ree et al. (2015) studies effect of doubling salaries on the intensive margin.
 - This paper studies effect of lowering salaries on the extensive margin.

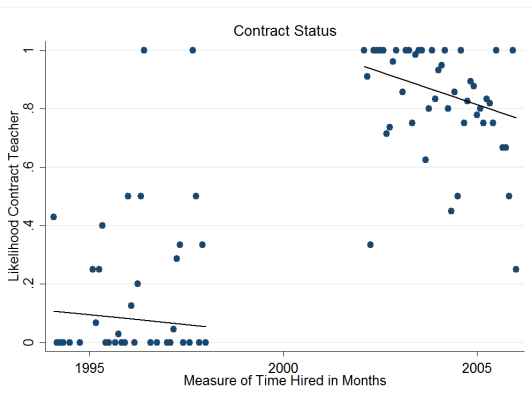
Effect of the Nuclear Test on Dollar Deposits in Pakistan



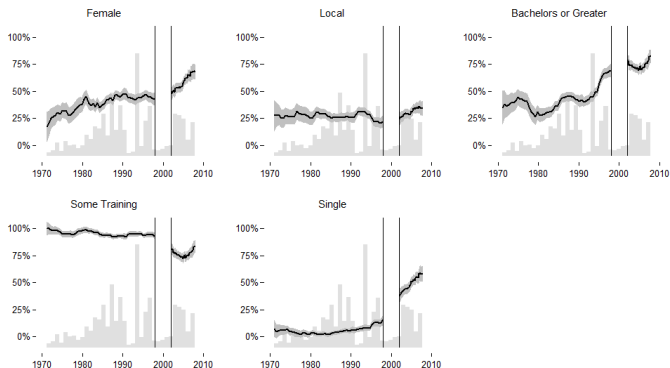
Contract Teacher Program (Cyan, 2009)

- The government moved dramatically from hiring almost all teachers on permanent contracts to hiring almost all teachers on temporary contracts.
- Teachers received 3-5 year contracts.
- Renewal was officially based on performance evaluation (in surveys, 45 percent of teachers agreed).
- No formal process for regularizing contract teachers.
 - 71 percent of teachers said job did not offer them an opportunity for “professional growth.”
 - 95 percent reported working on a temporary contract for more than 3 years.
- Hiring freeze from 1998 to 2001.

Effect of the Regime Change on Teacher Contracts



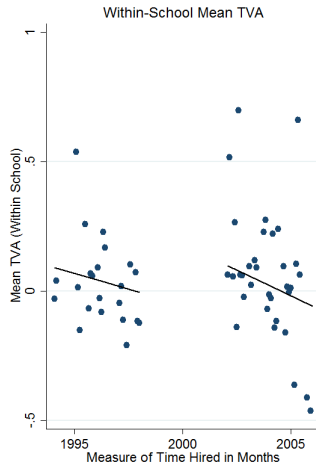
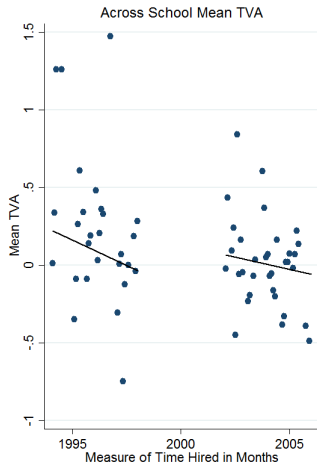
Effect on Teacher Characteristics



Teacher Characteristics Estimates

	(1) OLS	(2) SE	(3) RD (3 Year)	(4) SE	(5) RD (4 Year)	(6) SE
Log(Salary)	-0.284***	0.062	-0.554*	0.273	-0.444**	0.205
Bachelor's	0.318***	0.032	0.003	0.186	0.109	0.140
Some Training	0.003	0.031	0.013	0.120	0.010	0.096
Local	-0.017	0.037	-0.006	0.178	-0.066	0.134
Age Started	0.072***	0.024	1.193	1.550	0.943	1.116
Single	0.148***	0.032	-0.006	0.176	0.053	0.136
Female	-0.005	0.044	0.288	0.254	0.273	0.190
Mean Teacher English Score	0.326***	0.080	0.570**	0.231	0.319	0.248
Mean Teacher Urdu Score	0.076	0.067	0.429	0.342	0.217	0.336
Mean Teacher Math Score	-0.013	0.080	0.604	0.427	-0.502	0.375

Effect on TVA



Effect on TVA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Mean	SE	One-Sided	N	Within School	SE	One-Sided	N
	TVA		T-test		Mean TVA		T-test	
OLS (Full Sample)	-0.004	0.042	0.541	1,337	0.024	0.026	0.181	1,278
RD (Full Sample)	-0.004	0.052	0.533	1,337	0.056	0.041	0.088	1,278
RD (2 Year)	0.840	0.550	0.068	227	0.360	0.322	0.137	201
RD (3 Year)	0.219	0.241	0.184	376	0.254**	0.123	0.022	336
RD (4 Year)	0.350	0.234	0.070	393	0.193*	0.097	0.026	350
RD (5 Year)	-0.074	0.120	0.732	661	0.035	0.057	0.268	604
RD (6 Year)	-0.026	0.106	0.598	690	0.040	0.053	0.225	631
RD (7 Year)	-0.036	0.106	0.632	692	0.035	0.052	0.250	632

Is the Quality of Contract Teachers Declining Over Time?

- The policy change may affect the quality of the teaching pool with a lag.
- Therefore, we also want to see if the quality of incoming contract teachers is declining.
- Identification problem: Observe newer contract teachers with less experience.
- Solution: Differences-in-differences where we compare inexperienced contract teachers hired later and earlier to permanent teachers.
 - Result: No evidence quality is decreasing over time.
 - Coefficient: -0.007 (0.024).

Outline

- ① Data
- ② TVA Results
- ③ Regime Change Results
- ④ **Conclude**

Conclusion

- Teacher quality is important in low-income countries.
- As in the United States, besides experience, most observable teacher characteristics do not predict quality.
- Exception: content knowledge.
- Teacher salaries are not related to teacher quality in the public sector (but are in the private sector).
- Students of teachers hired on 35% lower salaries perform as well or better than students of permanent teachers.

LEAPS Testing Structure

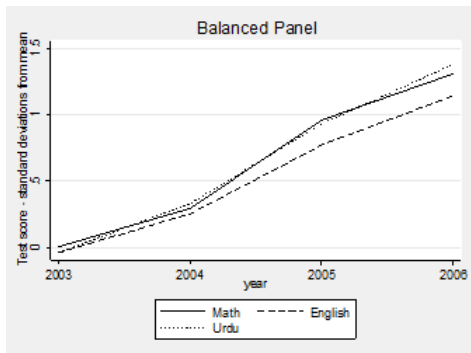
	(1) Number of Teachers	(2) Number of Students	(3) Teachers in Schools With > 1 Teacher With Tested Students	(4) Students in Schools With > 1 Teachers With Tested Students
Public, Rd 1	486	8,340	4	131
Private, Rd 1	303	3,617	0	0
Public, Rd 2	593	9,327	214	3,290
Private, Rd 2	336	3,340	97	846
Public, Rd 3	1007	16,946	884	15,320
Private, Rd 3	579	6,777	524	6,247
Public, Rd 4	1103	15,357	812	12,610
Private, Rd 4	599	5,911	478	5,020

▶ Back

Public School Students Used in TVA Estimation

Grade	Rounds		
	<u>Student-Years</u>		
	2	3	4
1	1	1	0
2	3	1	5
3	347	34	364
4	6,676	1,135	6,449
5	6	6,373	865
6	0	5	4,653
7	0	0	8

Learning Over Time

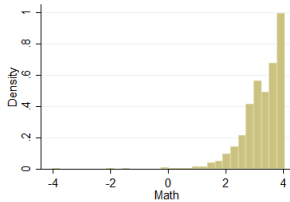
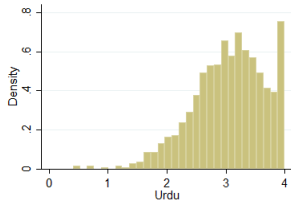
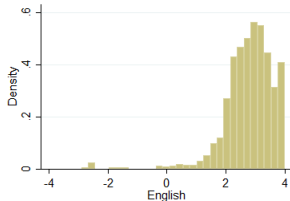


▶ Back

What Does a Test Score Mean?

	Year 1 Prop correct	Year 2 Prop correct	Year 3 Prop correct	Year 4 Prop correct
Total kids	6,038	6,038	6,038	6,038
English				
Eng 12: Match picture with word, Banana	0.631	0.75	0.834	0.873
Eng 18: Fill missing letter for picture, Cat	0.68	0.743	0.817	0.853
Eng 19: Fill missing letter for picture, Flag	0.287	0.299	0.478	0.554
Eng 30: Fill missing word in sentence	0.276	0.332	0.441	0.535
Eng 43: Construct sentence with word 'deep'	0.01	0.014	0.037	0.108
Eng 44: Construct sentence with word 'play'	0.024	0.027	0.113	0.218
	0.318	0.361	0.453	0.524
Math				
Math 1: Count number of moons, write number	0.622	0.687	0.797	0.749
Math 9: Add 3 + 4	0.903	0.91	0.951	0.94
Math 12: Multiply 4 x 5	0.603	0.641	0.759	0.811
Math 24: Add 36 + 61	0.855	0.878	0.922	0.93
Math 25: Add 678 + 923	0.561	0.595	0.712	0.745
Math 27: Subtract 98 - 55	0.698	0.756	0.826	0.856
Math 30: Multiply 32 x 4	0.522	0.569	0.703	0.756
Math 32: Divide 384 / 6	0.193	0.245	0.456	0.541
Math 34: Cost of necklace, simple algebra	0.092	0.148	0.257	0.278
Math 39: Convert 7/3 into mixed fractions	0.014	0.046	0.07	0.145
	0.5063	0.5475	0.6453	0.6751
Urdu				
Urdu 3: Match picture with word, Book	0.739	0.822	0.916	0.946
Urdu 4: Match picture with word, Banana	0.736	0.824	0.906	0.945
Urdu 5: Match picture with word, House	0.538	0.601	0.679	0.755
Urdu 10: Combine letters into word	0.737	0.792	0.861	0.897
Urdu 12: Combine letters into word	0.372	0.45	0.537	0.627
Urdu 19: Antonyms, Chouta	0.44	0.502	0.688	0.792
Urdu 20: Antonyms, Khushk	0.368	0.493	0.623	0.693
Urdu 36: Complete passage for grammar	0.293	0.391	0.563	0.678

Teacher Knowledge



▶ Back

Alternative Methods I: Empirical Bayes (Chetty et al., 2004; Kane and Staiger, 2008)

- Multiply noisy estimate of TVA (such as TVA generated by our method) by an estimate of its reliability.
- Estimate reliability as ratio of signal (TVA) variance to signal plus noise (student and year variance).
- Within classroom variance gives student variance.
- Covariance between average residual in teacher's class in t and $t - 1$ gives teacher variance.
- Variance of classroom component is the remainder of the residual's variance.

Alternative Methods I: Empirical Bayes (Chetty et al., 2004; Kane and Staiger, 2008)

Problems:

- Estimating teacher variance this way requires that a teacher's quality is time-invariant or stationary.
- To satisfy this assumption, we must include experience fixed effects.
- We cannot control for experience without subsuming the contract effect.
- Instead, teacher fixed effects capture mean teacher quality over the surveyed period, including mean experience effects.

Alternative Methods II: Child Fixed Effects (Rockoff, 2004)

- Method:
 - Include child fixed effects in the TVA estimating equation to further control for selection.
- Problem:
 - Relies on children switching teachers.
 - In Pakistan, teachers teach multiple grades, so this reduces the effective sample by 54 percent.
 - Mis-entered teacher ids may dominant the new sample, biasing estimates.

Alternative Methods II: Child Fixed Effects (Rockoff, 2004)

For example, assume:

- Students are identical and TVA is randomly distributed.
- A student has a probability $p = 0.1$ of changing teachers each year.
- An ID has a probability $e = 0.01$ of being incorrectly entered.

Then, there are three cases where a change appears to take place:

- Id was incorrectly entered and no change occurs: probability = $0.01 \times 0.9 = 0.009$
- Id is correctly entered and a change happens: probability = $0.99 \times 0.1 = .099$
- Id is incorrectly entered and a change occurred: probability = $0.1 \times 0.01 = 0.001$

So, the probability a teacher id is mis-attributed in the effective sample is $\frac{0.01}{(0.009+0.099+0.001)} = 0.09$

Alternative Methods II: Child Fixed Effects (Rockoff, 2004)

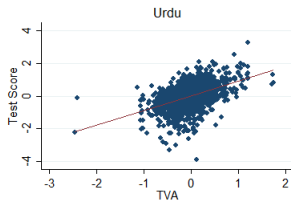
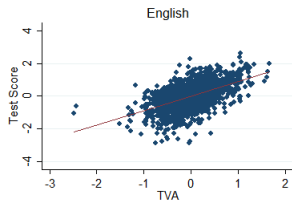
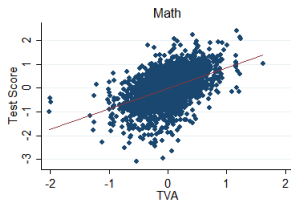
More generally, assume:

- Students are identical and TVA is randomly distributed.
- A student has a probability p of changing teachers each year.
- An ID has a probability e of being incorrectly entered.

Then,

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

Graphical Results



▶ Back

Sampling Error

$$E(\Phi) = E\left(\frac{\sigma^2}{\sum_{kt} N_{kt} \mathbf{1}_{k \neq j}}\right).$$

- ϕ is bias in the estimated covariance.
- N_{kt} is the number of students in the class of teacher k in year t
- σ^2 is the variance of idiosyncratic shocks at the student-level.

▶ Back