Online Appendix for "Can public rankings improve school performance? Evidence from a nationwide reform in Tanzania"

Jacobus Cilliers[†] Isaac Mbiti[‡] Andrew Zeitlin[§]

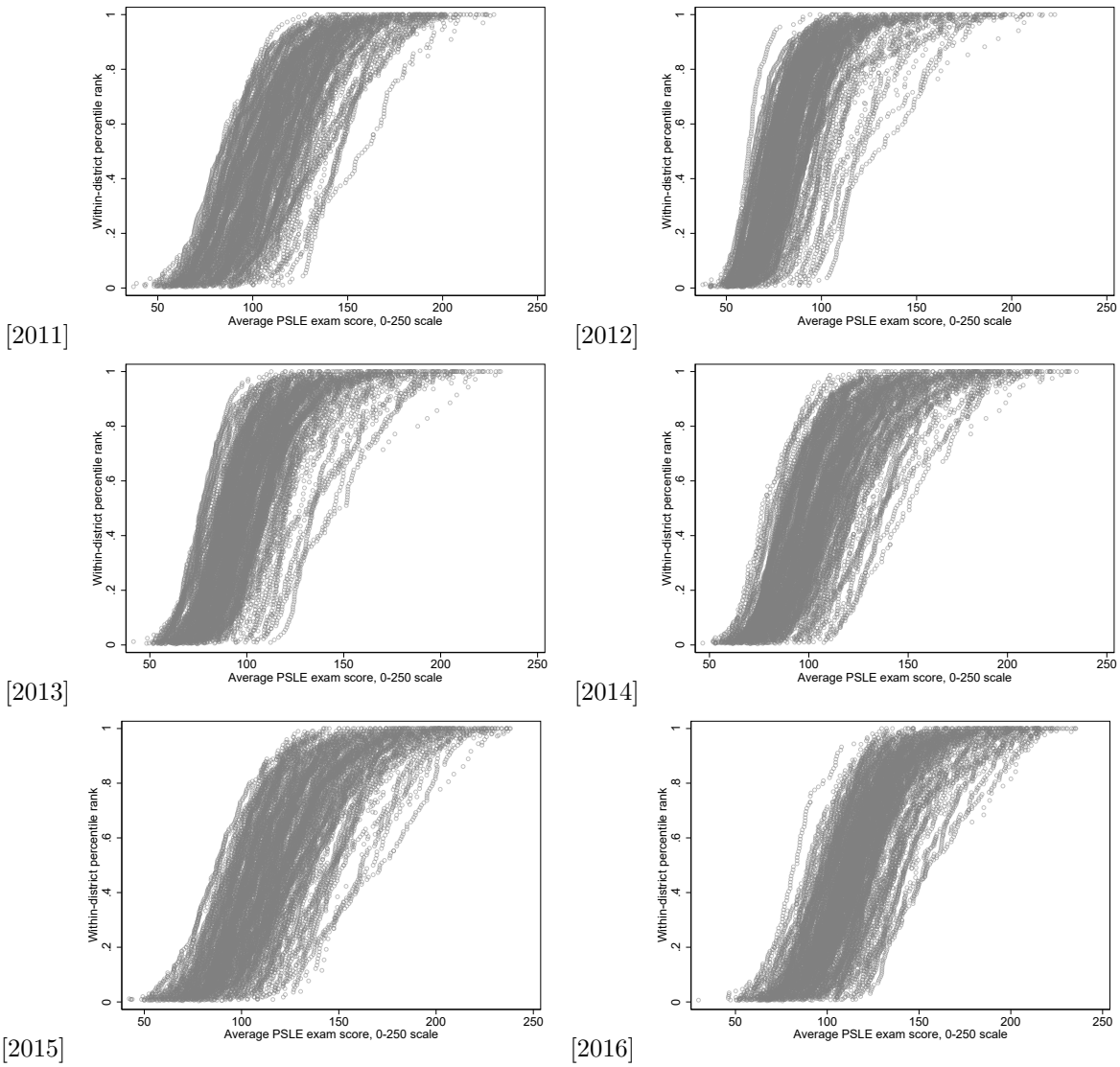
[†]Georgetown University; ejc93@georgetown.edu

[‡]University of Virginia, J-PAL, BREAD, NBER, IZA; imbiti@virginia.edu

[§]Georgetown University and Center for Global Development; az332@georgetown.edu

Appendix A Supplementary figures

Figure A.1: Absolute PSLE scores and within-district ranks, all years



Note: Figures illustrates relationship between the average PSLE score of students in a school and the school's rank in its district in the corresponding year.

Appendix B Supplmentary tables

Table B.1: Big Results Now in Education Initiatives

	I. Pressure to perform	
Official school ranking	Ranks all government primary (and secondary schools) by pass rates in PSLE (and CSEE). Each exam has 10 performance bands, which are classified as green, yellow, or red. Results publicly posted and widely disseminated. Both national rankings and district rankings were distributed.	Fully implemented through 2016.
School incentive scheme	Annual monetary and non-monetary incentives for primary and secondary schools that have most improved their performance in the national exams (PSLE & CSEE).	Partially implemented: 60 primary schools received financial awards in 2015. Almost 4000 non-monetary awards (certificates) distributed to primary and secondary schools starting in 2016. No awards prior to 2015.
	II. Teacher motivation	
Teacher motivation	Providing both non-monetary incentives (certificates) to high performing teachers, as well as clearing all outstanding payment arrears for teachers.	Partially implemented: outstanding claims reduced by a third by 2016.
	III. Back to basics	
National 3R Assessment	Early learning assessments (Grade 2) under the 3R (reading, writing, and arithmetic) assessment program on a set of randomly selected schools.	Implemented: assessment conducted in 2016.
3R Teacher Training	Teacher training program for Grade 1 and 2 teachers on how to teach reading, writing and arithmetic most effectively to this age group. Through a cascade model 37.5% of schools in 40 low-performing districts (out of 136) will be trained.	Implemented: Almost 60,000 teachers were trained by 2016. Training started in 2014.
Student Teacher Enrichment Programme (STEP)	STEP trained primary and secondary school teachers on how to identify and support low performing students. Teachers were trained on how to conduct diagnostic tests to determine which students need extra coaching, as well as how to develop curriculum and conduct classes for low performing students.	Partially implemented: teachers from 5500 primary schools were trained. Training started in 2014.
	IV. School management and finance	
School improvement toolkit	The programs aims to train head-teachers of primary and secondary schools on best practices in the management of schools. A practical toolkit of these practices was distributed to head-teachers.	Implemented: More than 16,000 primary schools received the materials. Distribution started in 2014.
Capitation grants	Timely disbursement of sufficient capitation grants for primary and secondary schools; as well as equalization of funding per student per district (about USD 4.6 per primary student and USD 11.6 per secondary student).	Late implementation: prior to 2016, 31% of schools received funds on time. In 2016 about 90% of schools received funds on time.

Notes: The delays in implementation were due to lack of funding. BRN was not adequately funded until donors provided funding in 2015. The information displayed in this table is compiled from a series of World Bank Project Implementation Status Reports and Tanzania Government reports (Government of Tanzania, 2015; World Bank, 2015, 2016b,c).

Table B.2: Pre-BRN outcome levels

	Number Passed	Average Marks	Exam sitters	Pass rate (%)
0–10th Percentile in 2012	8.4	72.8	47.6	15.5
	(15.3)	(15.1)	(34.2)	(17.0)
10–20th Percentile in 2012	11.0	75.6	54.1	17.9
	(16.9)	(14.2)	(36.7)	(16.5)

Notes: Table presents means and standard deviations of key outcomes in 2012. Calculated from NECTA data.

Table B.3: Impacts on the probability of a school being in the bottom decile or quintile in its district

	Bottom decile		Bottom quintile	
	(1)	(2)	(3)	(4)
0-10th percentile in previous year	-0.056***	-0.309***	-0.085***	-0.362***
	(0.021)	(0.018)	(0.025)	(0.021)
10-20th percentile in previous year	-0.007	0.006	-0.030	-0.275***
	(0.013)	(0.013)	(0.018)	(0.016)
Diff-diff	Yes	Yes	Yes	Yes
School fixed effects	No	Yes	No	Yes
Control lagged exam score	Yes	Yes	Yes	Yes
Observations	77731	77431	77731	77431
R-squared	0.089	0.391	0.126	0.428

Notes: Each column represents a separate regression. All specifications include district-by-year fixed effects, flexible controls for lagged test scores, and indicators pre-reform associations between district-rank deciles and subsequent outcomes. Reported coefficients correspond to the *differential* effect of being ranked in the associated decile of within-district performance in the post- (vs. pre-)reform period, compared to the middle six deciles. Columns (2) and (4) include school fixed effects. The outcome in columns (1) and (2) is an indicator for whether the school is in the bottom decile of its district rank; the outcome in columns (3) and (4) is an indicator for whether the school is in the bottom quintile of its district rank. Standard errors are clustered at the district level.

Table B.4: Testing for sorting

	(1) Exam sitters
0-10th percentile in previous year	-1.771**
	(0.759)
10-20th percentile in previous year	-2.126***
	(0.622)
80-90th percentile in previous year	1.318*
	(0.681)
90-100th percentile in previous year	1.451
	(0.895)
Ward: 0-10th pctile in previous year	-1.081
	(1.075)
Ward: 10-20th pctile in previous year	1.242
	(1.315)
Ward: 80-90th pctile in previous year	-1.553
	(1.288)
Ward: 90-100th pctile in previous year	1.011
	(1.159)
Observations	77431
R^2	0.910

Each column represents a separate regression. All specifications include district-by-year fixed effects, flexible controls for lagged test scores, and indicators pre-reform associations between district-rank deciles and subsequent outcomes. The first four coefficients correspond to the differential effect of being ranked in the associated decile of within-district performance in the post (vs. pre-)reform period, compared to the middle six deciles. Coefficients marked 'Ward' correspond to the effect of the share of schools in the ward within a given decile, compared to the middle six deciles, in the post (vs. pre-) reform period. Standard errors are clustered at the district level.

Table B.5: Impacts of the reform on school exam performance (sample of SDI schools)

	(1) Marks	(2) Pass rate	(3) Number passed
0-10th percentile in previous year	10.00**	0.130***	4.413**
	(4.49)	(0.04)	(1.82)
10-20th percentile in previous year	-2.913	-0.0104	0.267
	(3.92)	(0.04)	(1.55)
Control mean, post BRN	114.32	0.64	56.25
Observations	1534	1534	1534
R-squared	0.819	0.830	0.953

Each column represents a separate regression, estimated with flexible controls for lagged test scores and district-by-year and school fixed effects. Coefficients correspond to the effect of being ranked in the corresponding decile of within-district performance in the post-reform period, compared to the middle six deciles. The sample is schools in the SDI sample, for the years 2013 to 2016. Standard errors are clustered at the district level.

Table B.6: Impacts on student learning outcomes using SDI survey data

	(1) Math	(2) English	(3) kiSwahili
0-10th percentile in previous year	-0.027	0.377	-0.382
	(0.279)	(0.292)	(0.712)
10-20th percentile in previous year	0.034	0.430	0.281
	(0.195)	(0.310)	(0.453)
Post-BRN mean: 0-80th percentile	0.02	0.02	0.00
Observations	760	760	760

Each column represents a separate regression, estimated with flexible controls for lagged test scores and district-by-year and school fixed effects. Coefficients correspond to the effect of being ranked in the corresponding decile of within-district performance in the post-reform period, compared to the middle six deciles. Student learning outcomes are calculated as the first principal component of exam responses and standardized to have mean zero and variance one. The sample is schools in the SDI sample, for the years 2013 and 2016. Standard errors are clustered at the district level.

Table B.7: Within-district ranking impacts on enrollment, by gender

	Grades 4-6		Grade 7	
	(1)	(2)	(3)	(4)
	Male	Female	Male	Female
0-10th percentile in previous year	-0.753	-0.020	-0.805*	-0.841*
	(1.130)	(0.829)	(0.411)	(0.430)
10-20th percentile in previous year	1.056	0.873	-0.304	-0.175
	(0.726)	(0.688)	(0.277)	(0.296)
Control mean	95.20	101.93	24.14	27.41
Observations	31150	31150	31150	31150
R-squared	0.960	0.967	0.910	0.912

Table replicates results of Table 3, disaggregating enrollment outcomes by gender. Data are EMIS data, available for 2015–2016. Each column represents a separate regression, estimated with flexible controls for lagged test scores and district-by-year and school fixed effects. Coefficients correspond to the effect of being ranked in the corresponding decile of within-district performance in the post-reform period, compared to the middle six deciles. In columns (1) and (2), the outcome is total enrollment in grades 4–6, and in columns (3) and (4) the outcome is enrollment in grade 7. Standard errors are clustered at the district level.

Table B.8: Within-district ranking impacts, by experience of negative shocks in preceding year

	10 pctile drop		15 pctile drop		20 pctile drop		25 pctile drop	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Drop	None	Drop	Drop	Drop	None	Drop	None
0-10th percentile in previous year	-0.578	3.528***	-5.073	3.093***	6.825	3.019***	-18.558	2.940***
	(3.084)	(0.511)	(11.286)	(0.496)	(5.217)	(0.492)	(13.720)	(0.482)
10-20th percentile in previous year	0.055	1.253***	5.031	1.066***	4.382	0.969***	-3.379	0.917***
	(1.643)	(0.358)	(5.125)	(0.360)	(9.894)	(0.357)	(33.090)	(0.344)
Observations	9466	49428	3338	52380	1875	54694	932	56463
R^2	0.799	0.837	0.827	0.829	0.849	0.822	0.864	0.815

Table replicates the main results for average PSLE scores reported in Table 2, for subsets of schools that have (odd columns) or have not (even columns) experienced a negative shock of the corresponding strength in the preceding year. Each column represents a separate regression, estimated with flexible controls for lagged test scores and district-by-year and school fixed effects. Reported coefficients correspond to the *differential* effect of being ranked in the corresponding decile of within-district performance in the post- (vs. pre-)reform period, compared to the middle six deciles. Standard errors are clustered at the district level.

Table B.9: Single difference estimates on PSLE performance, by school size

	Average score			Pass rate		
	(1)	(2)	(3)	(4)	(5)	(6)
	All schools	Smallest	Large	All schools	Smallest	Large
0-10th percentile in previous year	5.648***	5.522***	5.734***	0.058***	0.047***	0.061***
	(0.682)	(1.249)	(0.739)	(0.008)	(0.015)	(0.009)
10-20th percentile in previous year	2.025***	1.840**	2.088***	0.020***	0.012	0.021***
	(0.414)	(0.891)	(0.452)	(0.004)	(0.010)	(0.005)
Observations	62792	12181	50611	62792	12181	50611
R^2	0.601	0.619	0.605	0.528	0.506	0.545

Each column represents a separate regression, estimated using data from post-reform years only. All specifications include flexible controls for lagged test scores and district-by-year fixed effects. Reported coefficients refer to the effect of being ranked in the associated decile of within-district performance, compared to the middle six deciles. In Columns (1)–(3) the outcome is the average PSLE score (ranging from 0–250); in Columns (4)–(6) it is the pass rate. Columns (2) and (5) use only schools in the bottom quintile of the school size distribution (based on number of testtakers in 2012); columns (3) and (6) use the complementary set of large schools. Standard errors are clustered at the district level.