

Scaling up What Works: Experimental Evidence on External Validity in Kenyan Education

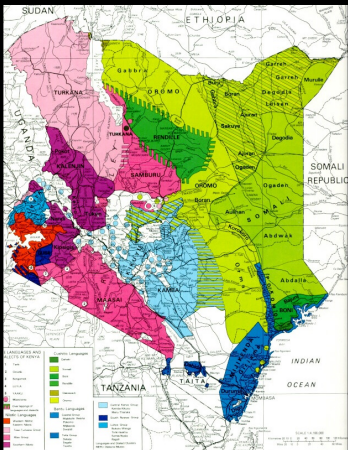
Tessa Bold	Goethe University & IIES
Mwangi Kimenyi	Brookings Institution
Germano Mwabu	University of Nairobi
Alice Ng'ang'a	Strathmore University
Justin Sandefur	Center for Global Development

May 9, 2013

Contract teachers

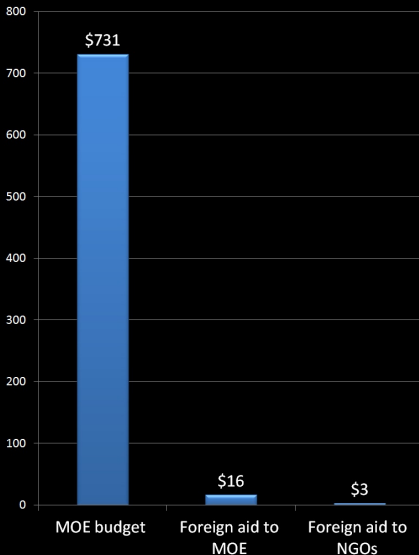
- ▶ Muralidharan & Sundararaman (2008)
Andhra Pradesh
Contract teachers \Rightarrow +0.15 std. dev.
- ▶ Duflo, Dupas, & Kremer (2009)
Western Kenya
Contract teachers \Rightarrow +0.21 std. dev.
Class size reduction \Rightarrow no effect on scores

Geography



Institutions

Millions of USD per annum spent on primary education



Scale per se



Average TSC Salary Sh.19,400 ≈ \$260 / month

 Sh.10,000 ≈ \$135 / month

Average PTA Salary Sh.4,200 ≈ \$56 / month

Scale per se



Average TSC Salary	Sh.19,400	≈	\$260 / month
	Sh.10,000	≈	\$135 / month
Average PTA Salary	Sh.4,200	≈	\$56 / month

Outline

Experimental design & context

Institutions

Horse race

Mechanisms

Geography

Conclusion

Outline

Experimental design & context

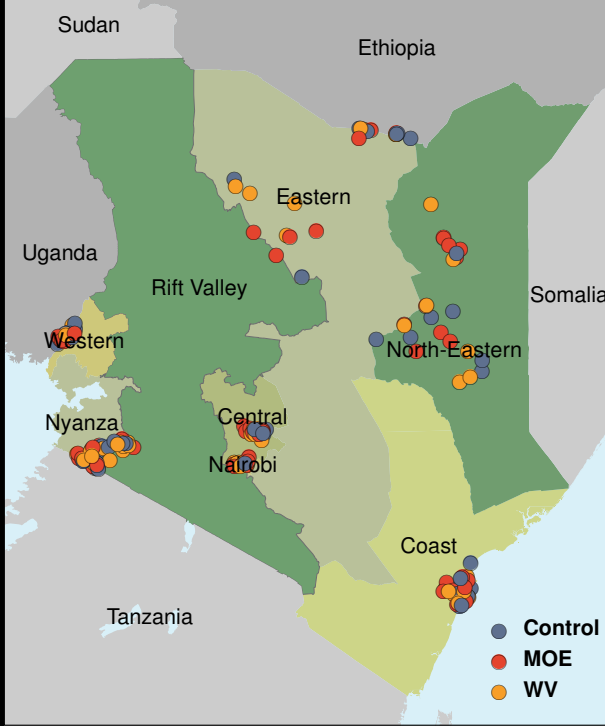
Institutions

Horse race

Mechanisms

Geography

Conclusion



Experimental Design

▶ Sampling

- ▶ All 8 provinces, 14 (non-random) districts
- ▶ Random sampling of schools w/ PTR > median

School-level randomization

- ▶ 192 schools
- ▶ 64 NGO, 64 Gov, 64 control

Intervention

- ▶ 1 add'l teacher per school
- ▶ Assigned to grade 2 or 3 in 2010
- ▶ 17 months exposure, immediate follow-up testing

Cross-cuts

- ▶ SMC training
- ▶ Central/local hiring
- ▶ High/low salary

Experimental Design

▶ Sampling

- ▶ All 8 provinces, 14 (non-random) districts
- ▶ Random sampling of schools w/ PTR > median

School-level randomization

- ▶ 192 schools
- ▶ 64 NGO, 64 Gov, 64 control

Intervention

- ▶ 1 add'l teacher per school
- ▶ Assigned to grade 2 or 3 in 2010
- ▶ 17 months exposure, immediate follow-up testing

Cross-cuts

- ▶ SMC training
- ▶ Central/local hiring
- ▶ High/low salary

Experimental Design

▶ Sampling

- ▶ All 8 provinces, 14 (non-random) districts
- ▶ Random sampling of schools w/ PTR > median

School-level randomization

- ▶ 192 schools
- ▶ 64 NGO, 64 Gov, 64 control

Intervention

- ▶ 1 add'l teacher per school
- ▶ Assigned to grade 2 or 3 in 2010
- ▶ 17 months exposure, immediate follow-up testing

Cross-cuts

- ▶ SMC training
- ▶ Central/local hiring
- ▶ High/low salary

Experimental Design

▶ Sampling

- ▶ All 8 provinces, 14 (non-random) districts
- ▶ Random sampling of schools w/ PTR > median

School-level randomization

- ▶ 192 schools
- ▶ 64 NGO, 64 Gov, 64 control

Intervention

- ▶ 1 add'l teacher per school
- ▶ Assigned to grade 2 or 3 in 2010
- ▶ 17 months exposure, immediate follow-up testing

Cross-cuts

- ▶ SMC training
- ▶ Central/local hiring
- ▶ High/low salary

Experimental Design

▶ Sampling

- ▶ All 8 provinces, 14 (non-random) districts
- ▶ Random sampling of schools w/ PTR > median

School-level randomization

- ▶ 192 schools
- ▶ 64 NGO, 64 Gov, 64 control

Intervention

- ▶ 1 add'l teacher per school
- ▶ Assigned to grade 2 or 3 in 2010
- ▶ 17 months exposure, immediate follow-up testing

Cross-cuts

- ▶ SMC training
- ▶ Central/local hiring
- ▶ High/low salary

Project Timeline

Jul 2009 Baseline evaluation for pilot

Aug 2009 Union lawsuit

Jun 2010 Pilot teachers placed in schools (NGO & Gov)

Oct 2010 Gov hires 18,000 contract teachers

Sep 2011 18,000 made permanent

Oct 2011 Final evaluation of pilot

Project Timeline

- Jul 2009 Baseline evaluation for pilot
- Aug 2009 Union lawsuit
- Jun 2010 Pilot teachers placed in schools (NGO & Gov)
- Oct 2010 Gov hires 18,000 contract teachers
- Sep 2011 18,000 made permanent
- Oct 2011 Final evaluation of pilot

Outline

Experimental design & context

Institutions

Horse race

Mechanisms

Geography

Conclusion

Outline

Experimental design & context

Institutions

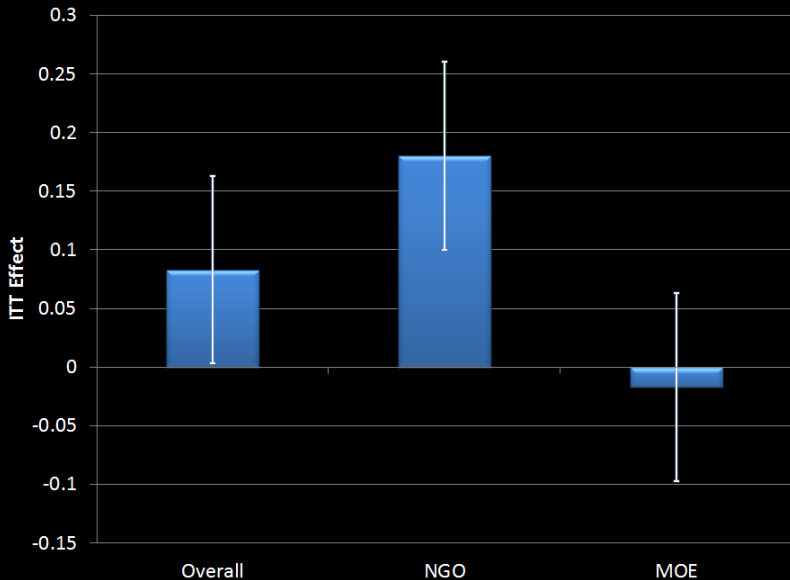
Horse race

Mechanisms

Geography

Conclusion

Treatment Effect of Contract Teachers on Test Scores



Experimental effects on teacher recruitment

Table: Labor supply of contract teachers

	(1)	(2)	(3)
Const.	.745 (.034) ^{***}	.686 (.047) ^{***}	.587 (.064) ^{***}
NGO implementation		.122 (.067) [*]	.123 (.065) [*]
High salary			.116 (.064) [*]
Local recruitment			.143 (.065) ^{**}
Obs.	2,044	2,044	2,044

Treatment Effects

Table: $Y_{ijt} = \alpha_j + \beta Z_{jt} + \gamma(Z_{jt} \times Gov_{jt}) + \delta_t + \varepsilon_{ijt}$

	ITT	LATE
Pooled:		
Z	.083 (.076)	
T		.119 (.108)
NGO vs Gov:		
Z	.180 (.084)**	
$Z \times Gov$	-.197 (.085)**	
T		.245 (.114)**
$T \times Gov$		-.270 (.122)**
Obs.	14,975	14,975

Outline

Experimental design & context

Institutions

Horse race

Mechanisms

Geography

Conclusion

Mechanisms (1 of 2)

	Gov.	NGO	Difference	Corr. with value added
	(1)	(2)	(3)	(4)
Teacher characteristics				
Female	.379	.203	.177 (.075)**	-.011 (.092)
Post-secondary education	.138	.014	.124 (.045)***	-.131 (.149)
Advanced prof. qualification	.069	.095	-.026 (.043)	.050 (.149)
Local institutions				
Friend/relative of teacher	.667	.373	.294 (.100)***	.051 (.100)
Presence	.628	.727	-.099 (.110)	.101 (.134)
Monitoring visit	.850	.961	-.111 (.053)**	.184 (.155)
National politics				
Ave. salary delay (months)	3.000	2.117	.883 (.291)***	-.056 (.034)*
Union represented me	.377	.149	.228 (.089)**	-.197 (.110)*
Took union action	.533	.471	.063 (.096)	-.068 (.097)

Mechanisms (2 of 2)

	Union identification		Test-score gains	
	(1)	(2)	(3)	(4)
$Z \times \text{Gov}$	0.084 (0.101)	0.157 (0.116)	-0.065 (0.149)	-0.075 (0.119)
$Z \times \text{NGO} \times \text{Union exposure}$	0.083 (0.120)		0.040 (0.183)	
$Z \times \text{Gov} \times \text{Union exposure}$	0.548*** (0.168)		-0.304* (0.154)	
$Z \times \text{NGO} \times \text{Exposure to gov't scale-up}$		-0.009 (0.115)		0.016 (0.143)
$Z \times \text{Gov} \times \text{Exposure to gov't scale-up}$		0.121 (0.154)		-0.258* (0.141)
Observations	100	95	102	107

Outline

Experimental design & context

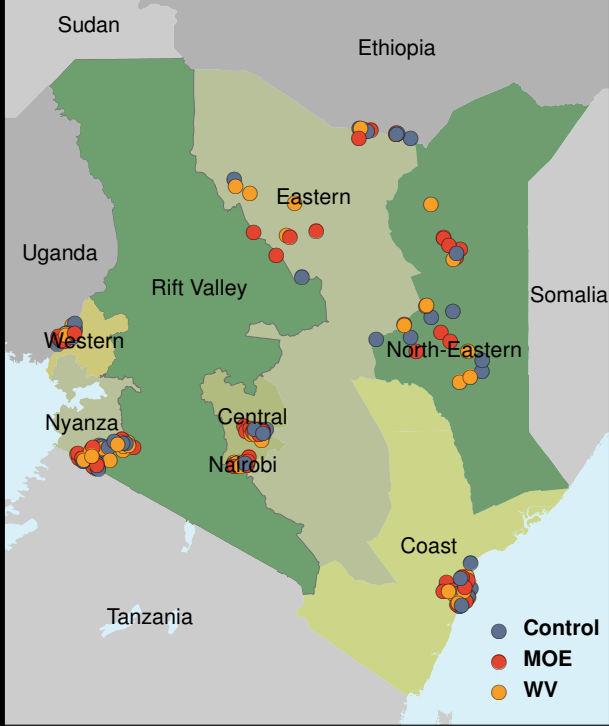
Institutions

Horse race

Mechanisms

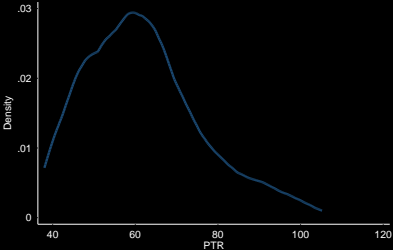
Geography

Conclusion

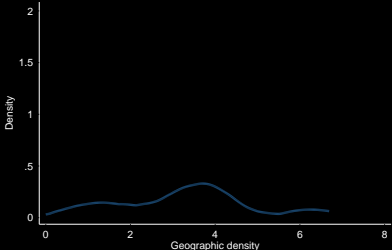


Heterogeneity

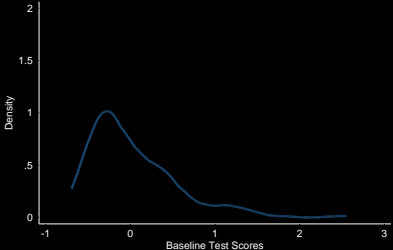
PTR



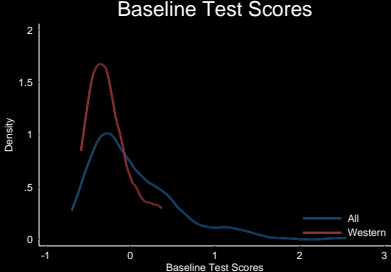
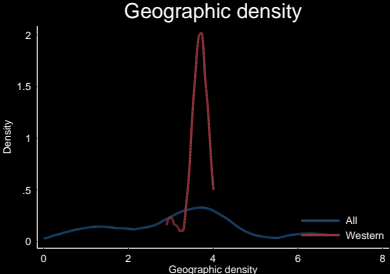
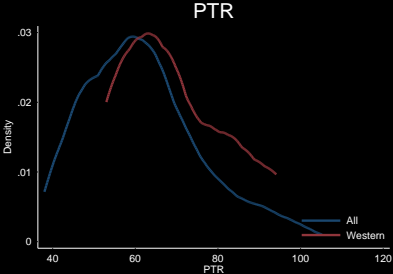
Geographic density



Baseline Test Scores



Heterogeneity



Heterogeneous treatment effects

Does impact vary across following dimensions? (overall, and for Gov and NGO individually)

- ▶ Geographic remoteness
- ▶ Initial pupil-teacher ratio
- ▶ Initial test scores

Heterogeneous treatment effects

Does impact vary across following dimensions? (overall, and for Gov and NGO individually)

- ▶ Geographic remoteness ✗
- ▶ Initial pupil-teacher ratio ✗
- ▶ Initial test scores (–) only in Gov sample

Heterogeneous treatment effects

Does impact vary across following dimensions? (overall, and for Gov and NGO individually)

- ▶ Geographic remoteness ✗
- ▶ Initial pupil-teacher ratio ✗
- ▶ Initial test scores (–) only in Gov sample

Western baseline scores 1/2 S.D. below mean

⇒ Gov-NGO gap 0.05 S.D. narrower in Western

Outline

Experimental design & context

Institutions

Horse race

Mechanisms

Geography

Conclusion

Conclusions (1 of 2)

- ▶ Geography & heterogeneous response
 - ▶ Intervention is progressive
 - ▶ But little reason question external validity from Western Kenya
- ▶ Institutions & partner selection bias
 - ▶ Horse race results: Institutions matter
 - ▶ e.g., local nepotism in gov't sector \perp of scale
- ▶ Scale & see-saw effects
 - ▶ Hint that gov't failure was a function of scale
 - ▶ e.g., union affiliation, salary delays

Conclusions (2 of 2)

- ▶ Lessons for impact evaluation
 - ▶ Is critique of external validity externally valid?
 - ▶ External validity vs. construct validity
 - ▶ Problem of IE not RCTs
 - ▶ NGOs as a laboratory vs. an accountability system

Compliance & Contamination

Table:

	All Schools		
	Treated	Control	Diff.
Compliance			
Class size	60.229 (3.179) ^{***}	69.047 (5.919) ^{***}	-8.818 (6.131)
Teacher ever in correct class	.953 (.020) ^{***}		
Teacher always in correct class	.729 (.043) ^{***}		
Contamination			
Log enrollment in treatment cohort	4.954 (.064) ^{***}	5.036 (.074) ^{***}	-.082 (.103)
Change in log cohort enrollment	-.109 (.023) ^{***}	-.093 (.035) ^{***}	-.016 (.040)
No. of teachers from 18,000 program	.667 (.107) ^{***}	.500 (.135) ^{***}	.167 (.177)

Compliance & Contamination

Table:

	Treated Schools		
	MOE	NGO	Diff.
Compliance			
Class size	60.470 (5.001)***	59.980 (3.687)***	.490 (6.131)
Teacher ever in correct class	.966 (.024)***	.938 (.035)***	.029 (.042)
Teacher always in correct class	.763 (.058)***	.688 (.072)***	.075 (.092)
Contamination			
Log enrollment in treatment cohort	4.951 (.070)***	4.957 (.105)***	-.007 (.094)
Change in log cohort enrollment	-.137 (.028)***	-.079 (.035)**	-.059 (.037)
No. of teachers from 18,000 program	.727 (.163)***	.607 (.140)***	.175 (.189)