ONLINE APPENDIX: MATERIAL NOT INTENDED FOR PUBLICATION

Contents:

Appendix A: Technocratic Selection Implementation Script

Enumerator A SCRIPT: Project Challenge and Manager Selection

STEP 1: Explain project challenge

READ TO GROUP: The Local Councils in Bombali and Bonthe are running a new exciting project challenge competition in your area. They are asking communities to submit proposals for small scale infrastructure (like construction of a latrine or drying floor, or repairs to a local school building). The Councillors will evaluate the proposals and pick the 20 best proposals as the winners. These 20 winning communities will receive **14 Million Leones** to use for implementing their projects. This is a lot of money! Your community is eligible to participate and I would like to encourage you to apply.

[HOLD UP THE PROPOSAL FORM FOR ALL TO SEE] This is the proposal form you will need to fill out to enter the project competition. I want this community to do well in this competition so will explain the things you need to put into a proposal and ask you to think about people in this community who would be good at putting these things together.

First, a strong project proposal needs a clear **description of the project**. This section tells the Council what the project will be, why the project solves an important problem or addresses an urgent need, and who will benefit from the project. To develop this description, you need a project leader who is good at identifying problems, coming up with solutions, making a persuasive argument ("sabi tok"), and who can read and write well.

Second, a winning project proposal needs to have a clear and reasonable **budget**. The budget lists all the items you will need to construct the project, how much they will cost, and where you will get them. It needs to show that your project will deliver value for money. You need a project leader who is familiar with these kinds of construction projects, knows where to get things, and how to get them at a good price, and someone who is good with numbers.

Third, a strong project proposal sets out a clear plan of work and **timeline**. This part of the proposal tells the Council who will do what and when. It should show that you know how to get things done: you can mobilize the workers you need, or know how to find a good contractor to work for you. You need a project leader who can set deadlines for each part of the project and get things done on time.

Before we leave today we will give you this project application form that you can use to submit the proposal. We will also tell you the date before which you need to submit the proposals. The proposals should be submitted in person to the District Council office in Makeni/Mattru Jong.

The winners announcement will be done in January 2017. You will receive an invitation to participate in the awards ceremony. We hope you will apply!

READ TO GROUP: Now I would like all of you to think about people in this village who are good at doing the things needed to develop a strong project proposal. I will step away from the group and let you think and talk about who would be good for this important job. We all know that the village headman has lots of experience running projects in this community. I would like you to also give me the names of 5 other people (in addition to the headman) that have these skills: they can **read and write**, they can come up with a **persuasive plan**, they know how to put a **budget** together, they are good at setting a **timeline**, **meeting deadlines** and **getting things done**. I will step away now so please call to me to come back when you have come up with the 5 people plus the headman.

STEP 3: Observe the proceedings

Step away outside the circle of the focus group and observe what happens.
Enumerator A: Fill out <u>TALLY SHEET A</u> below.
Enumerator B: Fill out <u>TALLY SHEET B</u> below.

STEP 4: Collect names of nominees / volunteers

Enumerator A: *When the community has finished its deliberation, rejoin the focus group and ask them to give you the names of the people they recommend.*

Name of Headman:_____

Name of 1st nominee:_____

Name of 2nd nominee:_____

Name of 3rd nominee:_____

Name of 4th nominee:_____

Name of 5th nominee:_____

NOTE: *if fewer than 5 nominees (in addition to the headman) were identified, only give the tests to the individual(s) selected by the focus group. If more than 5 nominees (in addition to the headman) were identified, ask the participants to rank the individuals and only work with the top 5 (plus the headman).*

READ TO GROUP: Thank you for these nominations. I would like to now ask each of these 5 nominated people to complete a short survey with me in private. The survey includes a test to measure the skills we talked about that are important for leading the project proposal: writing, making a project plan, doing a budget, working with numbers. The test will be done in private and the results will not be made public. Once all the tests are done, we will come back together as a group and I will unlock the project leader lottery. This lottery will randomly pick who will be the project proposal leader: it will tell us whether the leader for this project challenge competition will be A) the person with the highest score on the management test; or B) the village headman. I myself do not know which person the lottery will pick, and I cannot unlock the lottery until everyone completes the test. So let us please take a break and come back together at [TIME] to unlock the lottery and see who will lead

the project challenge competition for this village!

Complete the management tests with all 6 people above. Score the tests on site IN PRIVATE. When finished, see which person of the 5 NON-HEADMAN nominees had the highest score on the test. Make sure you know this person's name so you can announce it to the group if the lottery picks the HIGHEST SCORER to be the project leader. Do NOT share any information on how people scored on the management test.

STEP 6: RECONVENE THE FOCUS GROUP TO UNLOCK THE LOTTERY

READ TO GROUP: Thank you for coming back together. We can now unlock the project leader lottery! Remember, it will randomly pick whether A) the person with the highest score on the management test or B) the village headman will be the leader for the project challenge competition.

[UNLOCK THE LOTTERY: HOLD THE SCREEN UP SO THAT EVERYONE CAN SEE THE LOTTERY RUNNING. ANNOUNCE THE LOTTERY RESULT TO THE GROUP]

STEP 7: NEXT STEPS VARY BY LOTTERY RESULT

➔ **IF THE LOTTERY SAYS "HEADMAN LEADER":**

Explain that the lottery has randomly chosen the HEADMAN to be in charge of the project proposal for the challenge competition. Show the group the project application form and say that you are writing the HEADMAN down as the project proposal leader. Write his name on the application in front of the group. Walk over to the HEADMAN and give him the project application form. Explain that the proposal should be submitted in person by himself. Also give him the transportation voucher and explain that this can be redeemed when the proposal is submitted. Tell him that you hope he will put together a proposal for this village and that he will submit it to the Local Council.

Announce that the proposal needs to be submitted to [LOCAL COUNCIL ADDRESS] before the deadline [DATE]. Encourage them to apply.

Thank everyone for their time and wish them good luck with the project challenge competition!

END MEETING HERE AND GO TO VILLAGE INSPECTION SES SURVEY SECTION N

➔ **IF THE LOTTERY SAYS "HIGHEST SCORER":**

Explain that the lottery has randomly chosen the person with the highest management test score to be in charge of the project proposal for the challenge competition. Remind the group that you have used some tests to measure the skills needed for a strong proposal—reading

and writing, budget and costing, previous project experience—and that the tests have identified [NAME OF HIGHEST SCORER] as the person with the strongest skills for this particular opportunity. Show the group the project application form and say that you are writing [NAME OF HIGHEST SCORER] down as the project proposal leader. Write his name on the application in front of the group. Walk over to [NAME OF HIGHEST SCORER] and give him/her the project application form. Explain that the proposal should be submitted in person by the [NAME OF HIGHEST SCORER]. Also give him/her the transportation voucher and explain that this can be redeemed when the proposal is submitted. Tell him/her that you hope he/she will put together a proposal for this village and submit it to the Local Council.

Announce that the proposal needs to be submitted to [LOCAL COUNCIL ADDRESS] before the deadline [DATE]. Encourage them to apply.

Thank everyone for their time and wish them good luck with the project challenge competition!

   END MEETING HERE AND GO TO VILLAGE INSPECTION SES SURVEY SECTION N

   ➔ **IF THE LOTTERY SAYS "HIGHEST SCORER + TRAINING":**

Explain that the lottery has randomly chosen the person with the highest management test score to be in charge of the project proposal for the challenge competition. Remind the group that you have used some tests to measure the skills needed for a strong proposal—reading and writing, budget and costing, previous project experience—and that the tests have identified [NAME OF HIGHEST SCORER] as the person with the strongest skills for this particular opportunity. Show the group the project application form and say that you are writing [NAME OF HIGHEST SCORER] down as the project proposal leader. Write his name on the application in front of the group. Walk over to [NAME OF HIGHEST SCORER] and give him/her the project application form. Explain that the proposal should be submitted in person by the [NAME OF HIGHEST SCORER]. Also give him/her the transportation voucher and explain that this can be redeemed when the proposal is submitted. Tell him/her that you hope he/she will put together a proposal for this village and submit it to the Local Council.

Announce that the proposal needs to be submitted to [LOCAL COUNCIL ADDRESS] before the deadline [DATE].

READ TO GROUP: And, this village is very fortunate as you have qualified for a special one day training session that the Local Councils are offering in your area to teach you how to develop a successful project proposal. The session will cover the critical steps we discussed earlier: how to write a project description, how to draft a budget and how to set and meet deadlines, plus many other useful skills. I want to be sure that this village benefits from this training so will also cover the transport costs of [NAME OF HIGHEST SCORER] to participate in this important training.

Give [NAME OF HIGHEST SCORER] the TRAINING voucher that can be redeemed for full

transport costs plus food and drinks at the training.

Announce that the training session will be held at [LOCATION] on this day [DATE] at this time [TIME]. Encourage them to [NAME OF HIGHEST SCORER] to attend the training!

Thank everyone for their time and wish them good luck with the project challenge competition!

   END MEETING HERE AND GO TO VILLAGE INSPECTION SES SURVEY SECTION N

Appendix B. Expert Prior Elicitation Details

Before collecting and analyzing the data, we first established what experts in the field *thought* we would find. To do so, we fielded a survey among several distinct groups of experts: i) policymakers working for multilateral aid agencies (including the World Bank, the Department for International Development, the United Nations Development Programme and the International Rescue Committee) located mostly in OECD countries; ii) policymakers in Sierra Leone with knowledge of the GoBifo project; iii) economics graduate students in the United States (at University of California, Berkeley) and the Netherlands (at Wageningen University); iv) economics undergraduate students in Sierra Leone (at Fourah Bay College); and v) faculty directly involved in evaluating CDD projects (including the co-authors of this study) and other development economics researchers. This yielded 126 completed surveys in total, composed of 25 surveys from policymakers (12 in the OECD and 13 in Sierra Leone), 78 from students (17 undergraduate and 61 graduate students), and 23 from faculty. Survey response rates were quite high for all groups (e.g. 84% for faculty and 99% for graduate students) save the OECD policymakers (39% completion).

We organize our long-run study around the same twelve hypotheses that we focused on in our earlier paper. For each hypothesis, we asked experts to predict the point estimates we would find in the long-run, in standard deviation units, and also indicate their level of certainty for each prediction (following DellaVigna and Pope 2018, forthcoming). As in our earlier work, we then group these hypotheses and predictions into two main families, infrastructure and institutions, respectively.

There were two versions of the survey: the first provided detailed information on our medium run results and the second asked the respondent to make predictions without any information provided. We randomized which version was given to each expert, with a few exceptions (e.g. a small subset completed both versions).

Figure A1 shows mean predictions by type of expert for each of three outcomes (institutional change, infrastructure and entry into the grants competition). The solid horizontal solid line shows the point estimate in the long-run data, with the 95% confidence interval demarcated by the dashed lines. Table A1 shows mean predictions pooled across all experts for the proportion of communities that would enter the grants competition for each of the six treatment arms in Figure 1.

Figure A1: Expert Predictions of Long-run CDD Effects for All 5 Expert Types
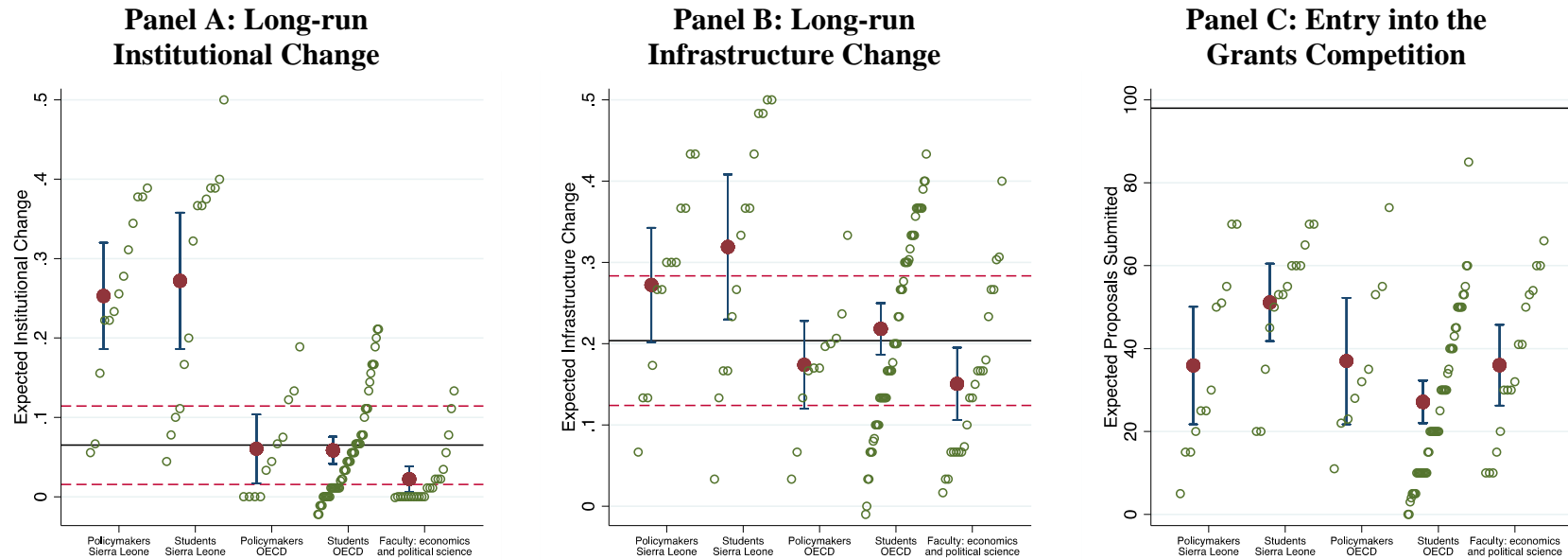
Notes: This figure presents predictions from 118 experts, including 71 economics students, collected during December 2016 and July 2017 before any data analysis. Panels A and B present expectations for CDD treatment effects measured in standard deviation units. The realized effect size is presented with solid black horizontal lines and the accompanying 95% confidence interval is demarcated by dashed horizontal lines. Panel C presents expectations about the percent of communities in the base case (no CDD, status quo chiefly control, or Arm 1 of Figure 1) that would enter the grants competition. The realized point estimates are: i) 0.066 standard deviation unit (standard error 0.025) CDD treatment effect on institutions for Panel A; b) 0.204 standard deviation unit (standard error 0.040) CDD treatment effect for infrastructure in Panel B; and c) 98.3% percent of communities entered the grants competition for Panel C. Expert predictions were closer to the realized value for the version of the survey that provided the short to medium run results for institutional change (p-value <0.01) but not statistically distinct for infrastructure (p-value=0.27).

Table A1: Predicted Entry into Grants Competition by Experimental Arms

2016 Assignment:	2005 Assignment:	
	CDD Control	CDD Treatment
Status Quo Chiefs	*Arm 1*	*Arm 4*
	35.5%	42.2%
	(23.0)	(21.1)
Technocratic Selection	*Arm 2*	*Arm 5*
	44.0%	53.9%
	(22.3)	(20.7)
Trained Technocrats	*Arm 3*	*Arm 6*
	53.6%	65.5%
	(23.5)	(20.9)
Realized entry, all communities:	98.3%	

Notes: This table presents mean expert predictions about the percent of communities that would enter the project challenge competition in each of the six distinct treatment arms in Figure 1. We pool predictions across all 118 experts, who were surveyed between December 2016 and July 2017, before data analysis.

Appendix C: Additional Specifications

Table A2: Lower Imputation Bound, Treatment Effects on Grants Competition Performance

	Proposal Score (index)	Technical Score	Expert Score	Gov't Score
	(1)	(2)	(3)	(4)
Panel A: Technocratic Selection versus CDD Institutional Reform				
Technocratic Selection	0.362**	0.465**	0.354**	0.267
	(0.168)	(0.191)	(0.172)	(0.179)
CDD	0.132	0.073	0.125	0.199
	(0.175)	(0.192)	(0.188)	(0.184)
Technocratic Selection * CDD	0.051	-0.025	0.173	0.006
	(0.221)	(0.247)	(0.232)	(0.236)
Observations	236	236	236	236
F-statistic (on TS and TS*CDD)	6.24	6.68	7.53	2.64
p-value	0.002	0.002	0.001	0.073
Panel B: Technocratic Selection and Managerial Training				
Technocratic Selection	0.252*	0.352**	0.245*	0.158
	(0.148)	(0.166)	(0.147)	(0.160)
Training	0.366**	0.311*	0.469***	0.319*
	(0.148)	(0.174)	(0.142)	(0.164)
F-statistic (on TS and TR)	10.39	9.23	13.99	5.17
p-value	<0.001	<0.001	<0.001	0.006
Observations	236	236	236	236

*Notes: i) significance levels indicated by * p < 0.10, ** p <0.05, *** p < 0.01; ii) robust standard errors; iii) specifications in Panel A pool the technocratic selection and training arms together (see Appendix Table A4 for full interaction model) and include strata for geographic ward and two balancing variables (distance to road and community size) from the original randomization; iv) specifications in Panel B include the two balancing variables and strata for ward crossed with CDD assignment; v) outcomes in columns 2 to 4 are mean effects indices, expressed in standard deviation units, standardized with respect to the mean and standard deviation of control Arm 1 (Arms 1 and 4) in Figure 1 for Panel A (B) (see Kling, Liebman and Katz 2007); vi) missing scores for the 4 non-submitting communities are imputed at the lowest observed score in the data; vii) outcome in column 1 is an equally weighted index of those in columns 2 to 4; vii) the Training term in Panel B captures the additional effect of training beyond that of technocratic selection; ix) the F-statistic and associated p-value evaluate the hypothesis that the listed terms are jointly equal to zero; and x) the sample for all specifications includes all communities in Figure 1.*

Table A3: Upper Imputation Bound, Treatment Effects on Grants Competition Performance

	Proposal Score (index)	Technical Score	Expert Score	Gov't Score
	(1)	(2)	(3)	(4)
Panel A: Technocratic Selection versus CDD Institutional Reform				
Technocratic Selection	0.366**	0.486**	0.348**	0.265
	(0.169)	(0.196)	(0.172)	(0.181)
CDD	0.001	-0.072	0.003	0.071
	(0.183)	(0.207)	(0.192)	(0.191)
Technocratic Selection * CDD	0.123	0.053	0.24	0.077
	(0.223)	(0.254)	(0.232)	(0.239)
Observations	236	236	236	236
F-statistic (on TS and TS*CDD)	7.71	8.34	8.82	3.40
p-value	0.001	<0.001	<0.001	0.035
Panel B: Technocratic Selection and Managerial Training				
Technocratic Selection	0.324**	0.438***	0.306**	0.228
	(0.141)	(0.158)	(0.143)	(0.154)
Training	0.309**	0.252	0.415***	0.260*
	(0.134)	(0.155)	(0.132)	(0.156)
F-statistic (on TS and TR)	11.34	10.71	14.63	5.27
p-value	<0.001	<0.001	<0.001	0.006
Observations	236	236	236	236

*Notes: i) significance levels indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; ii) robust standard errors; iii) specifications in Panel A pool the technocratic selection and training arms together (see Appendix Table A4 for full interaction model) and include strata for geographic ward and two balancing variables (distance to road and community size) from the original randomization; iv) specifications in Panel B include the two balancing variables and strata for ward crossed with CDD assignment; v) outcomes in columns 2 to 4 are mean effects indices, expressed in standard deviation units, standardized with respect to the mean and standard deviation of control Arm 1 (Arms 1 and 4) in Figure 1 for Panel A (B) (see Kling, Liebman and Katz 2007); vi) missing scores for the 4 non-submitting communities are imputed at the highest observed score in the data; vii) outcome in column 1 is an equally weighted index of those in columns 2 to 4; vii) the Training term in Panel B captures the additional effect of training beyond that of technocratic selection; ix) the F-statistic and associated p-value evaluate the hypothesis that the listed terms are jointly equal to zero; and x) the sample for all specifications includes all communities in Figure 1.*

Table A4: Full Interaction Model

	Proposal Score (index)	Technical Score	Expert Score	Gov't Score	Won a Grant
	(1)	(2)	(3)	(4)	(5)
Technocratic Selection	0.312	0.430*	0.289	0.217	0.100
	(0.194)	(0.231)	(0.199)	(0.209)	(0.065)
Training	0.162	0.185	0.165	0.138	0.002
	(0.197)	(0.234)	(0.194)	(0.218)	(0.078)
CDD	0.057	-0.018	0.056	0.132	0.048
	(0.182)	(0.207)	(0.193)	(0.191)	(0.047)
Technocratic Selection * CDD	-0.076	-0.076	-0.058	-0.094	-0.070
	(0.267)	(0.307)	(0.273)	(0.287)	(0.088)
Training * CDD	0.349	0.192	0.564**	0.290	-0.057
	(0.255)	(0.308)	(0.253)	(0.292)	(0.097)
F-statistic (on TS, TR and interactions)	8.33	5.83	11.88	3.42	1.31
p-value	<0.001	<0.001	<0.001	0.010	0.266
Observations	236	236	236	236	236

Notes: i) significance levels indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; ii) specifications include strata for geographic ward and two balancing variables (distance to road and community size) from the original randomization; iii) robust standard errors; iv) outcomes coded to treatment arm mean for communities that did not submit a proposal in columns 2 to 5; v) outcomes in columns 2 to 4 are mean effects indices, expressed in standard deviation units, standardized with respect to the mean and standard deviation of Arm 1 in Figure 1 (see Kling, Liebman and Katz 2007); vi) outcomes in column 1 are an equally weighted index of those in columns 2 to 4; vii) outcomes in column 5 are expressed in proportions; viii) the F-statistic and associated p-value evaluate the hypothesis that the listed terms are jointly equal to zero; and ix) the sample for all specifications includes all communities in Figure 1 (Arms 1 to 6).

Table A5: Two-way Comparison of Technocratic Selection and CDD

	Proposal Score (index)	Technical Score	Expert Score	Gov't Score	Won a Grant
	(1)	(2)	(3)	(4)	(5)
Technocratic Selection	0.444***	0.534***	0.487***	0.312**	0.052
	(0.113)	(0.129)	(0.118)	(0.120)	(0.036)
CDD Treatment	0.123	-0.004	0.208*	0.167	-0.017
	(0.105)	(0.123)	(0.109)	(0.116)	(0.036)
Observations	236	236	236	236	236
F-statistic (on TS and CDD)	8.74	8.52	10.47	4.45	1.05
p-value	<0.001	0.003	<0.001	0.013	0.351

Notes: i) significance levels indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; ii) specifications include strata for geographic ward and two balancing variables (distance to road and community size) from the original randomization; iii) robust standard errors; iv) outcomes coded to treatment arm mean for communities that did not submit a proposal in columns 2 to 5; v) outcomes in columns 2 to 4 are mean effects indices, expressed in standard deviation units, standardized with respect to the mean and standard deviation of Arm 1 in Figure 1 (see Kling, Liebman and Katz 2007); vi) outcomes in column 1 are an equally weighted index of those in columns 2 to 4; vii) the F-statistic and associated p-value evaluate the hypothesis that the listed terms are jointly equal to zero; and viii) outcomes in column 5 are expressed in proportions.

Table A6: Technocratic Selection Effects for Simulated Winning Thresholds

	Winner, actual	Winner, 25th Percentile	Winner, 50th Percentile	Winner, 75th Percentile
	(1)	(2)	(3)	(4)
Technocratic Selection	0.067	0.101	0.113	0.088
	(0.044)	(0.069)	(0.075)	(0.066)
Training	-0.026	0.09	0.167**	0.051
	(0.048)	(0.064)	(0.076)	(0.072)
F-statistic (on TS and TR)	0.99	7.89	13.51	4.26
p-value	0.322	0.005	<0.001	0.040
Implied number of grants	20	178	120	61
Observations	236	236	236	236

*Notes: i) significance levels indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; ii) robust standard errors; iii) specifications include strata for geographic ward crossed with CDD assignment; iv) outcomes in column (2)-(4) are binary indicator for winning a grant at percentiles of the government proposal score distribution; and v) the F-statistic and associated p-value evaluate the hypothesis that the listed terms are jointly equal to zero.*

Table A7: Management Training and "Teaching to the Test"

	Panel A: "Copycat" measures				Panel B: Performance spillover measures		
	References sustainability	References multiple bids	References skills needed	Index	Says who will benefit	Says where items bought	Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Technocratic Selection	-0.100	0.231	0.329*	0.153	0.163	0.129	0.146
	(0.148)	(0.218)	(0.177)	(0.103)	(0.147)	(0.158)	(0.108)
Training	0.362**	-0.231	0.127	0.086	0.125	-0.320**	-0.097
	(0.179)	(0.225)	(0.198)	(0.114)	(0.127)	(0.153)	(0.103)
Constant	0.028	0.551	-0.032	0.182	0.192	-0.042	0.075
	(0.354)	(0.704)	(0.370)	(0.288)	(0.213)	(0.290)	(0.132)
Observations	236	236	236	236	236	236	236

*Notes i) significance levels indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; ii) specifications include fixed effects for geographic ward crossed with CDD assignment; iii) Panel A looks for evidence of "teaching to the test" by seeing whether trainees mechanically include reference in their proposals to topics covered by the training but not asked for on the application (e.g. the training emphasized the value of seeking multiple bids from contractors during project construction, a good practice for winners to use during implementation but not something that the application required, and column 2 shows that trainees were no more likely to include extraneous reference to it in their proposals); iv) Panel B takes the converse approach and evaluates whether the training had performance spillover effects on application questions that were not addressed in the training (e.g. the application asked for an explanation of who would benefit from the project, a topic not discussed during the training, and column 5 shows that trainees were no more conscientious in including explanation of who benefits in their proposal); and v) outcomes in columns 4 and 7 are summary indices for the multiple measures in each panel.*

Table A8: 2018 Infrastructure Assessment of Grant Competition Winners by Treatment Assignments

| | Mean, full sample | Technocratic Selection Experiment | | | | CDD Experiment | | | |
| | | Mean, technocrats | Mean, Status Quo Chiefs | Difference | p-value | CDD treatment | CDD control | Difference | p-value |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Is the infrastructure present and functional?	0.70	0.63	1.00	-0.38	0.16	0.89	0.55	0.34	0.11
Quality of construction (1=poor, 10=excellent)	6.80	6.56	7.75	-1.19	0.26	7.00	6.64	0.36	0.67
Total community financial contributions (US$)	218.3	173.8	396.5	-222.7	0.14	233.6	205.9	27.7	0.83
Infrastructure is located near chief's compound	0.40	0.38	0.50	-0.12	0.67	0.33	0.45	-0.12	0.61
Observations	20	16	4			9	11		

*Notes i) significance levels indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; ii) data is from the July 2018 field inspection of infrastructure projects that won an implementation grant from the government competition; and iii) estimates displayed are from two-sided t-tests for each of the two distinct experimental assignments.*

Table A9: Text Analysis of Proposal Content Across CDD Treatment

	Proposal mentions inclusiveness terms	Proposal mentions community institutions	Community Center project	Education project	Water project	Other project
	(1)	(2)	(3)	(4)	(5)	(6)
CDD	0.047	0.023	-0.115*	0.050	0.051*	0.014
	(0.064)	(0.063)	(0.060)	(0.037)	(0.031)	(0.064)
Constant	0.390**	0.298*	0.451***	0.146	0.162	0.242
	(0.168)	(0.161)	(0.167)	(0.123)	(0.111)	(0.153)
Observations	236	236	232	232	232	232

*Notes i) significance levels indicated by * p < 0.10, ** p <0.05, *** p < 0.01. ii) specifications include strata for geographic ward and two balancing variables (distance to road and community size) from the randomization; and iii) robust standard errors.*

Table A10: Long Run CDD Treatment Effects on Exact Panel Outcomes

	Treatment effect 2016	Naïve p-value	FDR adjusted q-value	Treatment effect 2009	Change over time
	(1)	(2)	(3)	(4)	(1) - (4)
Panel A: Institutions Family					
All outcomes in family (N=56)	0.064**	0.010	0.006	0.086**	0.037
	(0.027)			(0.030)	(0.028)
Collective action	0.104*	0.050	0.234	0.072	0.086
	(0.053)			(0.046)	(0.061)
Inclusion	0.034	0.351	0.539	0.084*	0.031
	(0.036)			(0.049)	(0.044)
Local authority	-0.032	0.573	0.632	0.110	-0.088
	(0.056)			(0.068)	(0.070)
Trust	0.107*	0.065	0.234	0.032	0.064
	(0.057)			(0.049)	(0.081)
Groups and networks	0.149**	0.038	0.234	0.056	0.121
	(0.071)			(0.045)	(0.074)
Access to information	-0.036	0.590	0.632	0.150*	-0.075
	(0.067)			(0.072)	(0.072)
Participation in governance	0.079	0.191	0.348	0.256**	-0.011
	(0.060)			(0.058)	(0.065)
Crime and conflict	-0.002	0.971	0.76	0.088	-0.012
	(0.063)			(0.062)	(0.074)
Political and social attitudes	0.154	0.216	0.348	-0.020	0.113
	(0.124)			(0.080)	(0.126)
Panel B: Infrastructure Family					
All outcomes in family (N = 29)	0.208***	<0.001	0.001	0.352**	-0.094***
	(0.041)			(0.035)	(0.036)
Project implementation	0.287***	<0.001	<0.001	0.875**	-0.450***
	(0.075)			(0.062)	(0.081)
Local public goods	0.228***	<0.001	<0.001	0.210**	0.024
	(0.046)			(0.041)	(0.041)
Economic welfare	0.240***	<0.001	<0.001	0.606**	-0.136**
	(0.056)			(0.061)	(0.062)
Observations	236			236	236

*Note: i) significance levels based on naive p-values and indicated by * p<0.10, ** p<0.05, ***p<0.01. ii) specifications include strata for geographic ward and two balancing variables (distance to road and community size) from the randomization; iii) robust standard errors; iv) all estimates are for hypothesis-level mean effects indices that equally weight component measures and are expressed in standard deviation units (see Kling, Liebman and Katz 2007); v) outcomes limited to those that were collected in the exact same fashion in both 2009 and 2016 survey rounds; and vi) 2009 data sourced from Casey et al (2012).*

Appendix D: CDD effects on Community Response to Ebola

Table A11: CDD Treatment Effects on Ebola Responsiveness

Outcome	Mean, controls	Treatment effect	Standard error	*p*-value
Mean Effects Index (all 13 indicators)	0.000	0.042	0.038	0.27
Community had an Ebola task force during the Ebola crisis	0.661	0.077	0.052	0.144
Correctly answers "No" to "Can Ebola spread through air?"	0.856	-0.005	0.040	0.896
Correctly answers "21" to "How many DAYS can it take for the first to symptoms arise?"	0.669	0.014	0.051	0.791
Total (of 11 possible) correct answers to questions about how one can get Ebola	5.220	0.006	0.187	0.974
Knows correct Ebola hotline number	1.000	0.000	0.000	.
Community created bye-laws in relation to Ebola	0.907	0.042**	0.019	0.029
Total (of 10 possible) correct answers regarding how to protect yourself against Ebola	4.975	-0.051	0.201	0.801
Correctly answers "No" to "Drinking salt water can help cure Ebola?"	0.958	0.030	0.019	0.112
Correctly answers "No" to "Drinking chloring can help cure Ebola?"	1.000	-0.009	0.009	0.319
Communities are more likely to go to formal health facilities (nurse, clinic)	0.924	0.014	0.030	0.631
Communities are more likely to go to formal health facilities for Ebola (nurse, clinic)	0.915	0.000	0.034	0.995
Correctly answers "No" to "Can someone spread Ebola even before they show symptoms?	0.695	0.030	0.052	0.564
Total correct answers (of 14 possible) regarding symptoms of Ebola	7.263	-0.230	0.232	0.323
Observations	236			

Note: i) significance levels based on naive p-values and indicated by * p<0.10, ** p<0.05, ***p<0.01. ii) specifications include strata for geographic ward and two balancing variables (distance to road and community size) from the randomization; iii) robust standard errors; and iv) this table includes 13 of 15 pre-specified primary outcomes in our PAP, excluding 2 outcomes that are observed for fewer than 20 communities in the

Appendix E: Heterogeneous effects of CDD with respect to community size

Anderson & Magruder (2017) reanalyze the data from Casey, Glennerster and Miguel (2012) with hybrid split sample-PAP econometric techniques and find that CDD was less effective in larger villages, as measured by the number of households. They find negative effects for public goods, access to information and participation in local governance in larger villages. The long-run data provides an opportunity for a fresh test. In Panel B of Table A12 we too find substantial heterogeneity for the infrastructure family: the point estimate on the interaction with village size (demeaned and scaled by 25 households) is -0.082 (standard error 0.038). The loss could reflect greater coordination challenges (Olson 1965) and/or lower per capita grants in larger communities. In Panel A, we do not find any interaction effect for the institutions family overall: the point estimate on the interaction is -0.033 *(*standard error 0.024). We do find negative, but only marginally significant, estimates for information and participation specifically.

Appendix Table A12: Heterogeneity in CDD Effects by Village Size

	Treatment effect in 2016	Standard error	Interaction w/ village size	Standard error
	(1)	(2)	(3)	(4)
Panel A: Institutions Family				
All outcomes in family	0.065***	(0.025)	-0.033	(0.024)
Collective action	0.098**	(0.050)	-0.038	(0.040)
Inclusion	0.033	(0.036)	0.016	(0.033)
Systems of authority	-0.032	(0.057)	0.008	(0.041)
Trust	0.107*	(0.057)	-0.044	(0.054)
Groups and networks	0.149**	(0.071)	-0.040	(0.052)
Information	-0.036	(0.066)	-0.144*	(0.073)
Participation	0.079	(0.060)	-0.114*	(0.067)
Crime and conflict	-0.002	(0.063)	0.040	(0.061)
Attitudes	0.154	(0.123)	-0.249**	(0.110)
Panel B: Infrastructure Family				
All outcomes in family	0.204***	(0.040)	-0.082**	(0.038)
Implementation	0.253***	(0.067)	-0.151**	(0.061)
Local Public Infrastructure	0.229***	(0.045)	-0.091**	(0.037)
Economic welfare	0.240***	(0.056)	-0.028	(0.051)
Observations	236			

*Notes: i) significance levels based on naive p-values and indicated by * p<0.10, ** p<0.05, ***p<0.01; ii) specifications include strata for geographic ward and two balancing variables (distance to road and community size) from the original randomization; iii) robust standard errors; iv) all estimates are for hypothesis-level mean effects indices that equally weight component measures and are expressed in standard deviation units (see Kling, Liebman and Katz 2007); and v) village size is demeaned by the average number of households per village (47) and divided by 25, so can be interpreted as the amount the outcome would change by if the community grew by 25 households.*

Appendix F: Raw Results for Long-run CDD Effects on Individual Outcomes

Appendix Table A13: Raw Results for CDD Effects on all Individual Outcomes

Row	Variable	Hypothesis	Mean, controls	Treatment effect	Standard error	Naïve p-value	N
H1: Implementation							
1	Does this community have a bank account?	H1, H3	0.042	0.240	0.040	0.000	236
2	Average score of all test takers	H1	41.681	1.145	1.500	0.445	235
3	Does this community have a Village or Community Development Committee (VDC or CDC)?	H1, H4, H10	0.432	0.173	0.057	0.003	236
4	Does this community have a village development plan (i.e. an agreed plan with specific priorities for what the community will do for its own development over the next few years)?	H1, H10	0.492	0.003	0.057	0.955	236
5	Has this community been visited by a Local Councillor in the past one year?	H1, H9	0.263	-0.074	0.046	0.109	236
6	Has this community been visited by a Ward Development Committee member in the past year?	H1, H9	0.102	0.019	0.035	0.579	236
H2: GoBifo improves the quality of local public services infrastructure.							
7	Ask the community: when was the last time this community brushed this foot path? [date]	H2, H4	-35.224	1.123	4.707	0.811	234
8	Does the community have a court barrie and is it functional?	H2	0.102	0.218	0.040	0.000	236
9	Does the community have a community center and is it functional?	H2	0.068	0.060	0.038	0.111	236
10	Does the community have a drying floor and is it functional?	H2	0.178	0.127	0.051	0.013	236
11	Does the community have a grain store and is it functional?	H2	0.119	0.198	0.051	0.000	236
12	Does the community have a latrine and is it functional?	H2	0.076	0.029	0.036	0.412	236
13	Does the community have a market and is it functional?	H2	0.000	0.025	0.013	0.064	236
14	Does the community have a palava hut and is it functional?	H2	0.042	0.019	0.028	0.487	236
15	Does the community have a public health unit and is it functional?	H2	0.110	-0.022	0.038	0.565	236
16	Does the community have a primary school and is it functional?	H2	0.466	0.125	0.058	0.030	236
17	Does the community have any wells (mechanical or bucket) and are any of them functional?	H2	0.661	0.000	0.057	0.997	236
18	Do any of the local sports teams have uniforms / vests?	H2	0.153	0.003	0.046	0.946	236
19	Does the community have a football / sports field and is it functional?	H2	0.619	0.160	0.054	0.003	236
20	Does the community have a traditional birth attendant (TBA) house and is it functional?	H2	0.025	0.124	0.032	0.000	236

Row	Variable	Hypothesis	Mean, controls	Treatment effect	Standard error	Naïve *p*-value	N
21	Ask to be taken to the nearest bush path. This should be a foot path (not a road for cars) that the community uses the most. Walk 100 steps down the path (i.e. look at the middle, not the start of the path). In your own opinion, how bushy is the path? [Answer indexed from 0 "very bushy" to 1 "very clear"]	H2, H4	2.653	-0.049	0.110	0.658	236
22	Since January 2006, has this community taken a project proposal to an external funder—like local government or NGO—for support? *Note that the community should have been the ones initiating the request.*	H2, H4	0.246	0.048	0.054	0.370	236
23	Does this community have a seed bank (i.e. where people can borrow rice or groundnuts to plant and repay after harvest)?	H2	0.085	0.049	0.040	0.225	236
H3: GoBifo improves general economic welfare							
24	Supervisor assessment that community is "much better off" or "a little better off" than other communities he/she has been to in this area	H3	0.364	0.091	0.058	0.114	236
25	When was the last time an outsider trader came to this village to buy agricultural or non-agricultural goods? (date - date of interview)	H3	-12.178	3.468	4.820	0.472	236
26	[From supervisor tour of community] Have you seen anybody selling packaged goods (cigarettes, crackers, etc) in this village today from their own home (i.e. not out of a store)?	H3	0.881	-0.015	0.040	0.705	236
27	Number of goods out of 10 common items (bread, soap, garri, country cloth/garra tie-dye, eggs/chickens, sheep/goats, palm oil/nut oil, coal, carpenter for hire/shop, tailor/dressmaker, blacksmith for hire/shop) that you can buy in this community today	H3	5.619	0.403	0.247	0.103	236
28	How many people have started a new business (even if it is small or informal) in this community in the past 2 years (since October 2007)? [Record name, type of business and year started]	H3	6.297	0.627	0.500	0.210	236
29	How many houses and small shops (including tables, boxes and kiosks) are selling packaged goods (like cigarettes, biscuits, etc) inside this community today?	H3	3.737	0.626	0.343	0.068	236
30	In the past 2 years (since October 2007), have you participated in any skills training (bookkeeping, soap-making), adult literacy (learn book) or vocation education courses (carpentry, etc.)?	H3	2.831	0.270	0.629	0.667	236
H4: GoBifo increases collective action and contribution to local public goods.							
31	Does this community have any communal farms?	H4	0.144	0.087	0.049	0.073	236
32	Does the primary school that children in the community attend have community	H4	0.746	0.066	0.049	0.179	236
33	Average quality of proposal as assessed by experts	H4	55.309	3.247	1.807	0.072	232
34	Do any people from different households here come together to sell agricultural goods or other petty trading as a group to markets outside of this village (i.e. heap the goods together and send one person to sell; NOT every person totes their own load)?	H4, H7, H8	0.347	-0.046	0.053	0.390	236

Row	Variable	Hypothesis	Mean, controls	Treatment effect	Standard error	Naïve p-value	N
35	Average quality of proposal as assessed by policy makers	H4	51.262	2.461	1.591	0.122	232
36	Average completeness of proposal	H4	10.026	-0.013	0.283	0.964	232
37	Whether the proposal is among the top 20 and a winner (as ranked by the Gobifo staff ar	H4	0.096	-0.011	0.037	0.767	232
38	Do any disabled people hold leadership positions in this community (like member of VDC, youth leaders, headman, women's leader, secret society head)?	H5	0.144	0.033	0.048	0.499	236
H5:GoBifo increases inclusion and participation in community planning and implementation, especially for poor and vulnerable groups; GoBifo norms spill over							
39	Did any disabled people (blind, polio, amputee, wheelchair, etc.) attend the last community meeting?	H5	0.398	0.102	0.063	0.103	236
40	In the past one year, have you attended any community meetings?	H5	-28.644	7.510	7.084	0.289	236
41	Enumerator record of total women (18+ years) present at gift choice meeting (field activity #1)	H5	2.449	-0.031	0.141	0.828	236
42	Enumerator record of total youths (18-35 years) present at gift choice meeting (field activity #1)	H5	2.288	-0.209	0.193	0.280	236
43	Did anyone take minutes (written record of what was said) at the most recent community meeting?	H5	0.220	0.075	0.056	0.181	236
44	Less concentrated deliberation in manager selection	H5	2.892	0.023	0.090	0.798	231
45	Less concentrated deliberation in manager selection	H5	1.416	0.013	0.057	0.813	192
46	Enumerator account of how democratically the group evenutally came to a decision about who the potential project managers ranging from 5 = open discussion followed by group vote to 1 = chief and/or elders decide without other input	H5	3.364	-0.002	0.094	0.982	235
47	Time of deliberation of manager selection process	H5	32.486	53.665	27.838	0.054	210
48	Enumerator record of total public speakers durings selection of potential project managers	H5	43.429	-2.772	2.584	0.283	213
49	Did a vote occur during the project leader nomination discussion	H5, H6	1.929	0.023	0.032	0.463	171
50	Record of total women (18+ years) in "important people" focus group list	H5	13.264	-0.570	1.176	0.628	216
51	Enumerator account of how actively women participated in the deliberation on the selection of potential project managers compared to men, ranging from 5 = no difference between women and men to 1 = women not active at all compared to men	H5	2.799	-0.122	0.132	0.356	232
52	Record of total youth (18-35 years) in "important people" focus group list	H5	6.009	-0.402	0.391	0.303	229
53	Enumerator account of how actively youth participated in the deliberation on the selectic	H5	3.035	0.173	0.153	0.259	229
54	Has this community had any problems with financial mismanagement/corruption in the past 2 years (since November 2014)?	H5	0.839	-0.020	0.044	0.656	236

Row	Variable	Hypothesis	Mean, controls	Treatment effect	Standard error	Naïve *p*-value	N
H6: GoBifo changes local systems of authority, including the roles and public perception of traditional leaders (chiefs) versus elected local government.							
55	How old is the current (or acting) village chief/ Headman?	H6	-59.301	-0.974	1.830	0.595	228
56	Enumerator reports on whether "chief decided" project leader nominations	H6	0.873	-0.050	0.043	0.241	235
57	Relative view of "do people in this community believe" Local Councilors as opposed to Chiefdom officials	H6	-0.119	-0.021	0.052	0.683	236
H7: GoBifo increases trust							
58	Are you a member of any credit or savings (osusu) groups?	H7, H8	2.432	0.476	0.285	0.095	236
59	In general, do people in this community believe the central government officials or do they think you need to be careful when dealing with them?	H7	0.314	0.013	0.051	0.794	236
60	In general, do people in this community believe chiefdom officials or do you have to be careful when dealing with them?	H7	0.195	0.053	0.048	0.272	236
61	In general, do people in this community believe Local Councillors or do you have to be careful when dealing with them?	H7	0.076	0.032	0.037	0.391	236
62	In general, do people in this community believe NGOs / donor projects or do you have to be careful when dealing with them?	H7	0.500	0.168	0.057	0.003	236
63	In general, do people in this community believe people from outside you own village / town / neighborhood or do you have to be careful when dealing with them?	H7	0.127	0.088	0.047	0.062	236
64	In general, do people in this community believe people from you own village / town / neighborhood or do you have to be careful when dealing with them?	H7	0.703	-0.069	0.057	0.224	236
H8: Gobifo builds and strengthens community groups and networks							
65	Are there any fishing groups / cooperatives in this community?	H8	0.246	0.037	0.042	0.380	236
66	How many active school PTA groups are there in this village?	H8	4.076	0.719	1.208	0.552	236
67	How many active religious groups (not just going to church/mosque) are there in this village?	H8	4.102	1.721	2.019	0.394	236
68	How many active groups for saving for special events (weddings, funerals) are there in this village?	H8	0.517	0.164	0.116	0.156	236
69	How many active seed multiplication groups are there in this village	H8	0.254	0.853	0.485	0.079	236
70	How many active social clubs are there in this village?	H8	1.441	0.183	0.164	0.264	236
71	How many active women's groups (general) are there in this village?	H8	0.983	-0.039	0.124	0.749	236
72	How many active youth groups (general) are there in this village?	H8	1.212	0.013	0.110	0.907	236

Appendix Table A13: Raw Results for CDD Effects on all Individual Outcomes (continued)

Row	Variable	Hypothesis	Mean, controls	Treatment effect	Standard error	Naïve *p*-value	N
H9: GoBifo increases access to information about local governance							
73	Supervisor assessment of whether there are any of the following items--awareness campaigns, financial information, development plan, minutes from any meetings, government policies, election information--visible anywhere around the village (i.e. on a notice board, school, clinic, shop, etc.)?	H9	0.117	0.005	0.018	0.805	236
74	Has this community been visited by the Paramount Chief in the past year?	H9	0.127	-0.023	0.040	0.561	236
H10: GoBifo increases public participation in local governance							
75	Did anyone in this community contest the party symbol in the 2008 local council elections?	H10	0.169	-0.006	0.044	0.899	236
76	Did anyone in this community stand for the most recent paramount chief elections?	H10	0.068	0.032	0.035	0.357	236
77	Did anyone in this community stand for the most recent section chief elections?	H10	0.280	0.016	0.057	0.777	236
78	Did anyone in this community stand for the most recent Ward Development Committee elections or get nominated for WDC?	H10	0.212	-0.011	0.048	0.813	236
H11: By increasing trust, GoBifo reduces crime and conflict in community.							
79	No conflict that respondent needed help from someone outside the household to resolve in the past one year	H11	-10.424	0.520	1.103	0.637	236
80	In the past 12 months, respondent has not been involved in any physical fighting	H11	-0.568	-0.124	0.270	0.646	236
81	In the past 12 months, no livestock, household items or money stolen from the	H11	-12.127	-1.406	1.267	0.267	236
82	During the last 12 months, respondent has not been a victim of witchcraft (juju)	H11	-1.441	0.441	0.351	0.208	236
H12: GoBifo changes political and social attitudes, making individuals more liberal towards women, more accepting of other ethnic groups and "strangers", and less tolerant of corruption and violence.							
83	Is the current (or acting) village chief/Headman a woman?	H12	0.034	-0.010	0.022	0.653	236
84	Is the current (or acting) village chief/Headman less than 35 years old?	H12	0.009	0.034	0.021	0.107	228

Notes: i) significance levels (per comparison p-value) indicated by * p <0.10, ** p <0.05, *** p <0.01; ii) specification that includes fixed effects for the disctrict council wards (the unit of stratification) and the two balancing variables from the randomization (total households and distance to road) with robust standard errors; iii) "per comparison" p values are appropriate for a priori interest in an individual outcome

Appendix G: Pre-analysis Plan

We include below the text of our pre-analysis plan with annotation to flag where the referenced specifications appear in the main text and appendix. The plan, with time stamps, can be found in the American Economic Association's registry for randomized control trials (https://www.socialscienceregistry.org/trials/1784), where detailed Excel sheets listing all outcome variables (referenced as "PAP Sheets 1, 2, 3 and 4") are also available for download.

Annotated to link
pre-specifications to
manuscript tables
30 August 2018

Pre-analysis Plan: Two Approaches to Community Development

10 March 2017

PIs: K. Casey, R. Glennerster, E. Miguel and M. Voors

Overview

This research project has four main components. The first evaluates the long run effects of a community driven development (CDD) program in Sierra Leone. The project devolved financial and implementation control over public services to communities, accompanied by intensive social facilitation. The second assesses a low cost technocratic alternative that identifies and supports high competence community members to take better advantage of development opportunities. It leverages local talent, addresses information barriers, and augments existing managerial capital with basic training in project management. A third component elicits expert beliefs about the efficacy of these two approaches and assesses their forecast levels and accuracy. A fourth line of inquiry examines whether participation in CDD affected community response to the Ebola crisis.

Registration timeline

We registered this study with the American Economic Association (AEA) Randomized Control Trial Registry on 16 November 2016. Our trial entry can be found here: http://www.socialscienceregistry.org/trials/1784. On 17 November 2016, we uploaded a data management plan that outlines who would have access to data when, and commits all PIs to not access any data with identifying information until after this PAP is lodged. Fieldwork commenced on 18 November 2016. Our Field Manager Angelica Eguiguren at IPA Sierra Leone was the only person who had access to the data at all times. She uploaded the data to a secure server and will invite the PIs to that dropbox as soon as the PAP is lodged. We lodged an email confirming PI adherence to the data management plan on 9 March 2017. We lodged this PAP on 10 March 2017. We have received IRB clearance from Stanford (#38846), the Government of Sierra Leone, Office of the Sierra Leone Ethics and Scientific Review Committee (3-11-2016, Wageningen (18-11-2016), Berkeley (2016099099) and MIT (#1612798296) for this trial.

Part I: Long run effects of CDD

Component Overview: Community Driven Development (CDD) is a participatory approach popular with foreign aid donors that involves communities directly in the financial management and implementation of local public goods. CDD has two main aims: i) improve the stock and quality of local public goods via the provision of block grants; and ii) democratize local decision-making via intensive social facilitation focused on the participation of marginalized groups.

In earlier work, we analyzed the medium run effects of the "GoBifo" CDD project in Sierra Leone (Casey, Glennerster and Miguel 2012).[1] GoBifo was implemented from 2005 to 2009 and provided roughly $5,000 in block grants and six months of dedicated social facilitation per community. The medium run study found substantial positive impacts on local public goods and economic activity, stronger links between the community and local government, and no evidence for more inclusive local decision-making.

---

[1] Casey K, Glennerster R, Miguel E (2012) Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan. Quarterly Journal of Economics 127 (4): 1755-1812.

During late 2016, we revisited the 236 communities in the original study to assess long term impacts. In the interim, 60 of the treatment communities received additional support from the GoBifo project. Specifically, these 60 communities received $1,300 for youth empowerment programs in 2010. We do not know how exactly the project management staff selected these 60 communities from the pool of 118 treatment communities, but it was not via random assignment.

Hypotheses: The 12 research hypotheses grouped into two families remain the same as those used in the earlier study.

- Family A of hardware outcomes: "GoBifo creates functional development committees" (H1); "Participation in GoBifo improves the quality of local public services infrastructure" (H2); and "Participation in GoBifo improves general economic welfare" (H3).

- Family B of software outcomes: "Participation in GoBifo increases collective action and contributions to local public goods" (H4); "GoBifo increases inclusion and participation in community planning and implementation, especially for poor and vulnerable groups; GoBifo norms spill over into other types of community decisions, making them more inclusive, transparent and accountable" (H5); "GoBifo changes local systems of authority, including the roles and public perception of traditional leaders (chiefs) versus elected local government" (H6);[2] "Participation in GoBifo increases trust" (H7); "Participation in GoBifo builds and strengthens community groups and networks" (H8); "Participation in GoBifo increases access to information about local governance" (H9); "GoBifo increases public participation in local governance" (H10); "By increasing trust, GoBifo reduces crime and conflict in the community" (H11); and "GoBifo changes political and social attitudes, making individuals more liberal towards women, more accepting of other ethnic groups and 'strangers', and less tolerant of corruption and violence" (H12).

Econometric Specifications: For Part I, the primary test of interest is evaluating long run effects of CDD at the family level. Our core specification evaluates treatment effects for Family A and B, using the following model:

$$Y_c^L = \beta_0 + \beta_1 T_c + X'_c \Gamma + W'_c \Pi + \varepsilon_c \qquad (1) \quad \boxed{\text{In Table 4}}$$

where Y_c^L is the mean index for each family for community c in the 2016 survey round; T_c is the GoBifo treatment indicator; X_c contains two village-level balancing variables from the randomization process (distance from a road and total number of households); W_c is a fixed effect for geographic ward, the administrative level on which the randomization was stratified; and ε_c is the usual idiosyncratic error term. The parameter of interest is β_1, the average long run treatment effect. We will construct mean effects indices following Kling, Liebman and Katz (2007).[3]

To interpret these effects, we will test whether long run effects differ from the medium run effects in areas where the medium run effects were nonzero (Family A). Here we will test for decay using the following model:

$$Y_c^L - Y_c^M = \gamma_0 + \gamma_1 T_c + X'_c \Lambda + W'_c \Theta + \mu_c \qquad (2) \quad \boxed{\text{In Table 4}}$$

---

[2] As before, that this is not an explicit objective of the GoBifo project leadership itself, but is a plausible research hypothesis.
[3] Kling, J., J. Lieberman and L. Katz (2007) Experimental Analysis of Neighborhood Effects, Econometrica, 75(1); 83–119

where the dependent variable is the difference in mean effects indices measured in the 2016 survey, Y_c^L, and 2009, Y_c^M. The coefficient of interest is γ_1, where $\gamma_1 < 0$ suggests that the treatment effect has dissipated over time for that hypothesis. A combination of failing to reject $\beta_1 = 0$ while rejecting $\gamma_1 \geq 0$ suggests that previously observed treatment effects have dissipated, while failing to reject $\beta_1 = 0$ and $\gamma_1 \geq 0$ presents a less conclusive middle ground that likely reflects greater noise in measuring long run outcomes and accompanying reductions in the power to detect treatment effects. Note that the exact set of outcomes varies between the 2009 and 2016 data collection rounds, so each index will incorporate the relevant outcomes for that particular survey round (see below).

The second test of interest is running Equations (1) and (2) at the hypothesis level where Equation (2) will again only be run for hypotheses with non-zero medium run effects.

Throughout our analysis, we will adjust for the fact that we are running more than one test on the same dataset by implementing false discovery rate (FDR) corrections. Research practice appears to be moving towards FDR and away from the more conservative familywise error rate (FWER) corrections where there are several tests of interest. Since our earlier paper used FWER corrections, we will also report them here to maintain consistency, but note that the preferred specifications use FDR. These adjustments run across the two families (Family A and Family B) or 12 hypotheses (H1 – H12) as relevant. See Benjamini, Krieger and Yekutieli (2006) and Anderson (2008).[4] For all tests, we will also report the "naïve" or "per comparison" p-value.

Our third test of interest highlights a few individual outcome measures from a new structured community activity (SCA). Here we will test for long run effects of GoBifo on the managerial capital of community members and the quality of proposals submitted to a project challenge competition run by the local District Councils (discussed in greater detail below). These outcomes measure whether the learning-by-doing experience of participating in GoBifo translates into long run differences in ability to act collectively and take advantage of development opportunities. We will test them as part of our larger research framework under H1 and H4, respectively, but also highlight them on their own as they capture an important channel through which GoBifo could lead to long run changes.

To further interpret the family- and hypothesis-level results, we will also estimate Equation (1) at the level of individual outcome (adjusting for FDR across all outcomes under a given hypothesis). Note that this reporting of all individual outcomes is for illustrative and interpretation purposes only.

Measurement and survey instruments: See ["SES - Endline 2016"]. The main data collection instrument for the long run effects closely follows the community modules used in the 2009 survey. This includes a focus group discussion with local leaders and enumerator physical inspection of community amenities and market activity. Where possible, we have included a community-level analogue of household level indicators included in the 2009 survey. In addition to economic and social outcomes, we include measures of institutional outcomes using the new project challenge SCA. These are captured in several instruments ["Managerial capital test", "Manager selection tally sheet enumerator A and B", "Submission survey", "Submission form", "Technical scoring", "Policy Scoring", "Expert Scoring"]. We did not repeat the household level survey due to budget constraints.

---

[4] Benjamini, Y., A. Krieger, and D. Yekutieli (2006) Adaptive Linear Step-Up Procedures That Control the False Discovery Rate, Biometrika, 93: 491–507. Anderson, M (2008) 'Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedaian, Perry Preschool, and Early Training Projects,' Journal of the American Statistical Association, 103 (484): 1481–1495.

Outcomes: See ["PAP, sheet 1"]. The table maps each individual outcome to the hypothesis of interest. To facilitate comparison to our earlier work, the first several columns of this table reproduce exactly those in the Appendix J: Raw Results from the supplementary materials to the 2012 QJE article. The list of outcomes has evolved in a few key ways. First, the present data collection uses only community modules and does not conduct household visits. Thus, all household level outcomes (indicated by "HH" in column K "2009 survey level") are omitted. Where possible, we have included a community-level analogue in the current survey (see column O "Additional question 2016"). Second, we exclude almost all conditional outcomes (i.e. those that are contingent on having a specific good in the community) that are only observed for a subset of villages. Third, as part of our new SCA, we designed measures that mirror some of the process-oriented 2009 SCA outcomes (e.g. unobtrusively counting the number of women who participate in a community decision).

The Casey et al (2012) paper included 334 outcomes, excluding the conditional variables a total of 206 variables remain (see Table 2 in the paper). The 2016 survey round includes 101 outcomes. Table 1 displays the number of outcomes by hypothesis. In total, 96 outcomes exactly match across both rounds. As a robustness analysis, we rerun Equation (1) and Equation (2) for both survey rounds at the family level [Table A10] restricting the analysis to the 96 variables that appear in both 2009 and 2016 survey rounds.

Table 1. Non-conditional outcomes by Hypothesis

Hypothesis	2009	2016	Matching outcome in both rounds
Family A			
H1	7	6	5
H2	18	17	17
H3	15	7	7
Family B			
H4	15	10	6
H5	47	19	19
H6	25	4	4
H7	12	8	8
H8	15	9	9
H9	17	4	4
H10	18	9	9
H11	8	4	4
H12	9	4	4
Total	**206**	**101**	**96**

Heterogeneous Treatment Effects: We will test for heterogeneous treatment effects along the same eight community-level dimensions we used (and measured) in our earlier analysis (total households, war exposure, average schooling, distance to road, historical domestic slavery, district, ethnic fractionalization and chiefly authority). As an exploratory exercise, we will use an automated process (LASSO and BART) to identify other dimensions that are correlated with heterogeneous effects to mine the data in a principled way.

Part II: Managerial Capital

Component Overview: To evaluate a technocratic alternative to CDD's intensive social facilitation model, we overlaid a new randomized experiment across the GoBifo treatment arms. We will test whether i) a more technocratic approach to identifying project leaders with high managerial capital, and ii) the provision of training in project management fundamentals, improves community ability to active collectively and take advantage of a new development opportunity. Specifically, all communities had an opportunity to

enter a project challenge competition run by the local District Councils that awarded US$2,000 implementation grants to the twenty best project proposals. We block randomized 80 communities to a management selection treatment arm (*MS*); 78 to a management selection plus training arm (*MST*); and 80 to a control or status quo (*SQ*) mechanism that favors the village headmen.

These three treatment arms were implemented by the research team enumerators on the data collection visits to communities at the end of the focus group discussion. In all three arms, enumerators explained the project challenge opportunity and the skills needed to develop a strong proposal. They asked the group to deliberate and nominate five individuals, in addition to the village headman, who had these skills. These 6 individuals were then asked to take a management test, in private, which was scored on site by enumerators. The focus group was then reconvened and a public lottery (implemented on a tablet device) determined treatment assignment for the village. In the *status quo (SQ) arm*, the village headman was designated as the project proposal leader. His name was written on the standardized project application form and he was given a transportation voucher to redeem if/when he submitted a proposal to the relevant Local Council. In the *manager selection (MS) arm*, the enumerators announced who was the highest test performer (of the 5 non-chief nominees), and designated that person on the submission form and provided the transport voucher. The *manager selection plus training (MSTR) arm* followed the same format as *MS* but also announced that the relevant ward development committee (most local tier of elected government) would hold a one day management training as part of the project challenge competition. Enumerators provided the date and location of the training, informed the group that the travel costs of the designated project leader will be reimbursed, and encouraged the designated project leader to attend the training.

The training sessions for *MSTR* covered: i) identification of local development needs and designing projects to address them; ii) costing local materials and developing itemized budgets; and iii) time management and planning to meet deadlines. Note that measures of proposal quality capture both items covered in the training and those that were not, to evaluate the extent to which any observed training effects reflect "teaching to the test."

Hypotheses: We plan to evaluate the following hypotheses:

- There is underutilized managerial capital in villages (H-II.1)

- Leveraging underutilized managerial capital leads to greater ability to act collectively and take advantage of local development opportunities (H-II.2)

- Lack of management skills constrains the ability to take advantage of local development opportunities (H-II.3).

Econometric Specifications: Our primary tests of interest estimate:

$$P_c = \delta_0 + \delta_1 MS_c + \delta_2 TR_c + W'_c \Psi + \zeta_c \qquad (3)$$

In Table 1, Panel B

where outcome P (i.e. proposal quality, test score of project leader) is measured for community c; MS is an indicator variable equal to one for assignment to the manager selection process (*MS* and *MSTR* arms) and zero otherwise; TR is an indicator for assignment to training (*MSTR* arm); W_c is a stratification fixed effect for geographic wards; and ζ_c the idiosyncratic error term. Hypotheses H-II.1 and H-II.2 test $\delta_1 = 0$. Hypothesis H-II.3 tests $\delta_2 = 0$. Deviation: *Wc* is ward crossed with CDD assignment, see footnote 7

For Hypothesis H-II.1 we have only one outcome, the test score of the project proposal leader. For Hypotheses H-II.2 and H-II.3 we have four measures of proposal quality so our primary specification will

A32

be a mean effects index. We will also report estimates for the individual scores. As a robustness check, we will exclude quality assessments that involve any input from GoBifo staff (although note all proposals were blinded during the review).

Several additional analyses will aid in interpreting these results (see [PAP Sheet 2] for details). We will:

1. Explore the extent to which the training reflects "teaching to the test." Explore where the training appears most effective.
2. Validate the management test by correlating test scores with proposal quality and explore relative predictive of power of subsection scores.
3. Validate the extent to which the distinct manager selection treatment arms translated into differences in who actually managed the project proposal process.
4. Compare the tests scores of the non-headman nominees to those of village headmen.
5. Evaluate which characteristics correlate with managerial capital test scores (i.e. age, gender, education, management experience, leadership position, etc.).
6. Test for heterogeneous response to training by management test score.
7. Test for interaction effects between participation in GoBifo and the *MS* and *TR* terms in Equation 3, noting that these tests are likely underpowered.

Measurement and Survey Instruments: We used several instruments to implement and evaluate this new SCA, see ["Managerial capital test", "Manager selection tally sheet enumerator A and B", "Submission survey", "Submission form", "Technical scoring", "Policy Scoring", "Expert Scoring" and data from the transcripts of the training].

Outcomes: See ["PAP, Sheet 2"]

Part III: Expert Beliefs

Component Overview: There have now been several randomized control trials of CDD projects in different countries, most of which find some positive impacts on economic outcomes and little effect on institutions. A key unanswered question is whether experts—in academia and more importantly in policy—are updating their beliefs about how effective CDD projects are. This is important in light of the large amounts of foreign aid at stake ($85 billion spent on CDD in about two decades by the World Bank alone, according to Mansuri and Rao 2012), and whether the accumulation of evidence impacts the allocation of donor funds. We surveyed students, academic and policy experts to elicit their beliefs (following DellaVigna and Pope 2016) about the long run effects of the Sierra Leone CDD project and to forecast how well communities will perform in the new project competition.[5]

We fielded this survey among several distinct groups of experts: i) policy makers working for multilateral aid agencies (including the World Bank, DfID, UNDP and IRC); ii) policy makers in Sierra Leone with knowledge of the GoBifo project; iii) economics graduate students in the US (at UC Berkeley) and the Netherlands (at Wageningen University); iv) economics undergraduate students in Sierra Leone (Fourah Bay College), v) researchers directly involved in evaluating CDD projects other development (economics) researchers; and vi) the PIs of this study. There were two versions of the survey: version 1 provided detailed information on our medium run results and version 2 asked the respondent to make predictions without any

---

[5] DellaVigna, S. and D. Pope, "Predicting Experimental Results: Who Knows What?" NBER Working Paper No. 22566, August 2016. See also Humphreys, M., R. Sanchez de la Sierra and P. van der Windt (2016) Social Engineering in the Tropics: A Grassroots Democratization Experiment in Congo, working paper.

information provided. For the majority of respondents, we randomized whether they completed version 1 or 2. A small subset completed both versions.

Hypotheses:

- Estimated long run treatment effects are not the same as the average prior beliefs of surveyed experts (H-III.1) Appendix B

- Average prior beliefs and forecast accuracy differ across groups of experts (H-III.2) Appendix B

- Prior beliefs about long run effects of the GoBifo project are more optimistic (e.g. predict larger positive long run effects) amongst policy makers compared to researchers (H-III.3) Appendix B

- Predictions under version 1 of the survey (that contains information on the medium run effects) are more accurate than under version 2 (H-III.4) Appendix B

Econometric Specifications: For Hypothesis H-III.1, we will evaluate whether the average prior belief across all six groups of experts are statistically distinguishable from the estimated long run treatment effects by GoBifo family and hypothesis. For H-III.2 we will test whether mean predicted effect size by family varies across groups, and assess which estimate is closest to the observed long run effects. H-III.3 tests whether the mean prior of expert groups i and ii more optimistic (predict large positive effects) than that of groups v and vi, at the family level (one sided test). Tests of H-III.4 whether prior beliefs are more accurate in version 2 compared to version 1 across all six groups. For H-III.4 we will use all the data. As a robustness check we will drop data from the subset of respondents that completed both versions of the survey.

We will run several additional descriptive analyses. These include testing whether respondents who report higher confidence in their estimates, and greater familiarity with the 2012 study, are more accurate in their predictions. For the new SCA project challenge, we will impute several estimates—regarding GoBifo treatment effects, the efficacy of training, and the impact of technocratic manager selection—and compare their mean values and accuracy across expert respondent groups.[6]

Measurement and Survey Instruments: See ["Expert Priors Survey"]

Outcomes: See ["PAP, sheet 3"].

Part IV: Impacts on Ebola

Component Overview: The recent outbreak of Ebola Virus Disease (EVD) in West Africa is the largest ever recorded. The crisis resulted in over 4000 deaths in Sierra Leone alone (about 11000 in total). The two districts where GoBifo was implemented were differentially effected, Bombali saw 1050 suspected cases and 391 deaths, while Bonthe was much less hit, with 5 suspected cases and 5 deaths. In addition to Communities suffered directly due to fear, illness and loss of life, and indirectly due to travel and trade restrictions resulting from imposed quarantines. The Ebola crisis provided a huge stress on communities at social, political and economic levels. We analyze if participation in Gobifo put communities in a better position to implement preventative measures and collaborate with local government. We report two secondary outcomes (i) we separate impacts on knowledge and collective action, and (ii) we investigate if Gobifo villages reported different Ebola case-loads.

---

[6] We exclude the study PIs (group vi) from this comparison. While the PIs had no access to the data, we did learn through communication with the field team that the number of submitted proposals was very high.

Hypothesis: Our main hypothesis is that "Participation in GoBifo increased knowledge, collective action and investments in preventative measures during the Ebola crisis". Table A11

Econometric Specifications: same as Equation (1) above. Our dependent variable is a mean effects index of all Ebola related outcomes. As secondary outcomes, we assess impacts in a mean effects index for knowledge and collective action outcomes separately.

We asses outcomes for the whole sample and restrict our sample to Bombali, which saw many more Ebola cases than Bonthe making the collective action outcomes more relevant.

To further interpret the hypothesis-level results, we will also estimate Equation (1) at the level of individual outcome, adjusting for FDR across outcomes. Note that this reporting of all individual outcomes is for illustrative and interpretation purposes only.

Measurement and survey instruments: see ["SES - Endline 2016", module J and K].

Outcomes: See ["PAP, sheet 4"].