

Abramovsky, Laura; Augsburg, Britta; Oteiza, Francisco

Research Report

Sustainable total sanitation - Nigeria: Report on the second rapid assessment

IFS Report, No. R134

Provided in Cooperation with:

Institute for Fiscal Studies (IFS), London

Suggested Citation: Abramovsky, Laura; Augsburg, Britta; Oteiza, Francisco (2017) : Sustainable total sanitation - Nigeria: Report on the second rapid assessment, IFS Report, No. R134, Institute for Fiscal Studies (IFS), London,
<https://doi.org/10.1920/re.ifs.2017.0134>

This Version is available at:

<https://hdl.handle.net/10419/201775>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Sustainable total sanitation – Nigeria

Report on the second rapid assessment

Laura Abramovsky
Britta Augsburg
Francisco Oteiza

Sustainable Total Sanitation – Nigeria

Report on the Second Rapid Assessment

Laura Abramovsky, Britta Augsburg and Francisco Oteiza;
in collaboration with Indepth Precision Consulting, Nigeria and WaterAid
Copy-edited by Rachel Lumpkin

The Institute for Fiscal Studies

Published by

The Institute for Fiscal Studies

7 Ridgmount Street
London WC1E 7AE

Tel: +44 (0) 20-7291 4800
Fax: +44 (0) 20-7323 4780
Email: mailbox@ifs.org.uk
Website: <http://www.ifs.org.uk>

© The Institute for Fiscal Studies, October 2017

Preface

This report is part of the formal research component of WaterAid's "Sustainable Total Sanitation in Nigeria" (STS Nigeria) project, funded by the Bill & Melinda Gates Foundation. Additional funding from the Economic and Social Research Council (ESRC) Centre for Microeconomic Analysis of Public Policy (CPP) is also gratefully acknowledged. The views expressed in this report are, however, those of the authors and not necessarily those of the funders. Neither are the views expressed necessarily those of the other individuals or institutions mentioned here, including the Institute for Fiscal Studies (IFS), which has no corporate view.

Data were collected in collaboration with Indepth Precision Consult (IPC), based in Abuja, Nigeria. IPC bears no responsibility for the interpretation of the data in this report. All respondents agreed to participate in the surveys, and were assured of the confidentiality of any identifying information gathered. The University College London Ethics Review Board and the National Health Research Ethics Committee of Nigeria have approved this study.

The authors would like to thank Melanie Lührmann and Juan Pablo Rud for their helpful comments, and Sam Crossman for his excellent research assistance. Any errors and all views expressed are those of the authors.

Contents

Executive Summary	1
1. Introduction	2
2. CLTS and Research Design	4
2.1 The CLTS intervention	4
2.2 Data collection	5
2.3 Research design	5
2.4 Empirical strategy	7
2.5 Triggerability of villages and interpretation of coefficients	8
3. Impact Evaluation	10
3.1 Attrition	10
3.2 Impact of the CLTS intervention	11
3.3 Heterogeneous impacts	15
3.4 Improved sanitation	18
4. Programme Spillovers	20
4.1 Results	26
5. Conclusions and Further Research	31
References	34

Executive Summary

In this report, we analyse the impacts on sanitation uptake achieved by the Community Led Total Sanitation (CLTS) intervention, almost two years after its implementation in 2015, in the Nigerian states of Ekiti and Enugu. We report findings from an analysis of data collected in the second rapid assessment of households (RA2), carried out during April 2017, almost two years after implementation. We also discuss these results in contrast to the findings from the previous data collection wave, from December 2015 (RA1), around six to twelve months after implementation. Evidence from the cluster randomised controlled trial (RCT) shows the following.

- By RA2, CLTS households exhibit 4 percentage points (pp) lower prevalence of open defecation (OD) than control households.
- However, the small and positive CLTS treatment effects on toilet uptake over the whole sample (around 4pp) observed in the RA1, which were driven by an increase in uptake in treatment areas, cease to be detectable due to a relatively faster increase in toilet coverage in control areas by the time of the RA2.
- Results on toilet uptake show almost no variation by type of household by the time of the RA2. At RA1, CLTS treatment effects were higher (around 6–7pp), among households with female heads, among households whose heads had low levels of education, and among households with seniors. These higher effects are detectable only for the last group by RA2. Equally, households with children, households with no debts and households with low asset wealth, who also appeared to have reacted strongly to the CLTS intervention by RA1, are indistinguishable from the rest of our sample by RA2.
- We see no evidence of a change in the share of improved/unimproved toilets in the study sample because of the CLTS intervention.

In summary, most of the impacts observed six months after the intervention (RA1) become undetectable two years after it (RA2). This could be due to programme spillovers within communities that contained both CLTS and control households. Accounting for this in two different ways, we find the following.

- Estimated CLTS impacts on toilet uptake and functioning toilet uptake measured at RA1 are 7pp and 5pp, respectively, if spillovers are accounted for by a specific methodology, showing an increase in magnitude and significance.
- Estimated CLTS impacts on functioning toilet uptake become statistically significant and of a magnitude of 6pp by RA2, more than two years after the intervention took place.
- Future research will aim to shed light on how these spillovers operate, and on their implications for both research and policy design.

1. Introduction

In this report, we analyse the progress of WaterAid UK's Project 'Sustainable Total Sanitation Nigeria – implementation, learning, research, and influence on practice and policy' (STS Nigeria). This report is part of the project's Formal Research Component and is a follow-up to two previous reports (Abramovsky et al. 2015, 2016). We focus on the impacts on sanitation uptake achieved by the CLTS intervention, almost two years after its implementation in 2015, in the Nigerian states of Ekiti and Enugu. We report findings from an analysis of data collected in the second rapid assessment of households (RA2), carried out during April 2017. The data collection, data analysis and the compilation of this report have been carried out by the Institute for Fiscal Studies (IFS) in collaboration with Indepth Precision Consulting, and with input from WaterAid.

In the preceding 'Rapid Assessment Report' (Abramovsky et al. 2016), we showed that the CLTS intervention increased average sanitation uptake by 3pp or 4pp and reduced OD by 5pp from a baseline of 60% in communities targeted by CLTS. In the RA1, sanitation uptake was measured as the ownership or construction of a toilet, six to twelve months after implementation. In this report, we show that by RA2 (two years after implementation), these average impacts on sanitation cease to be detectable due to an increase in toilet coverage in control areas and to a slight fall in toilet construction rates in CLTS communities. However, the reduction in OD is still detectable and of similar magnitude, at 4pp. We explore CLTS impacts on a sample of the population identified in the previous report to be particularly reactive to the intervention, and to have low baseline levels of coverage. We find that among households with female heads, among households whose heads had low levels of education and among households with seniors, CLTS impacts by RA1 were of 6–7pp; in the latter case, these effects are also detectable two years after the intervention. Households with children, which at RA1 showed slightly higher programme impacts than the rest of the sample, do not appear to have significantly higher levels of toilet ownership by RA2. Similarly, households with no debts and households with low asset wealth, who also appeared to react strongly to CLTS by RA1, are indistinguishable from the rest of our sample by RA2. Finally, we study the pattern of toilets built as a result of CLTS, and we show that the evidence does not suggest any change in the share of improved toilets in our sample.

At RA2, most of the programme impacts observed in RA1 have become undetectable, due to strong growth in construction rates among control households. This increase in coverage in control areas could perhaps be explained by programme spillovers within larger communities and small towns that contained both treatment and control triggerable units. Accounting for this in two alternative ways, we find that the estimated impacts measured at both RA1 and RA2 increase in magnitude and significance. We estimate programme impacts in the range of 5–6pp for the ownership of functioning toilets and for ownership or construction of toilets. At baseline, only 36% of our sample owned or was constructing a toilet, so this represents an increase of 17% in baseline coverage rates by RA1, attributable to CLTS. In the case of improved sanitation, accounting for programme spillovers leads us to find statistically significant impacts of 5pp by RA2. This is an equivalent percentage increase of around 17%, considering that, at baseline, improved toilet coverage was of 30%.

The remaining report describes these findings in more detail. Before showing the analysis of the results, in Section 2 we briefly describe the CLTS intervention that was implemented in the study areas, as well as the methodology that underlies this report. In Section 3, we discuss the main impacts of the intervention on household toilet uptake, and we explore particularly its heterogeneous impacts. In Section 4, we explore the possibility of programme spillovers driving our results. We conclude in Section 5.

2. CLTS and Research Design

Drawing heavily on the first Rapid Assessment report, in this section we describe the CLTS intervention, the data collected and the research design. We present a summarised version of what is described more length in Abramovsky et al. (2016).

2.1 The CLTS intervention

In the study areas, CLTS implementation started in early 2015, just after the completion of the baseline survey. The implementation can be broken down into four distinct phases, which we describe briefly here.

1. *Planning.* Organising the next two phases, mobilisation and triggering, is a desk-based activity that can take about four hours.
2. *Mobilisation.* The CLTS triggering team visits communities to be triggered and talks with community leaders. The aim of this visit is to engage the leaders and to agree on a date and a time for triggering activities to take place. The date should be chosen so as to be suitable for the majority of community members to attend. For large communities, a single date will be set for the triggering of multiple clusters concurrently or consecutively. Sometimes it requires two to three visits to set a date for triggering. Each visit takes between one and two hours, excluding travel time.
3. *Triggering.* On the agreed date, at least four staff members – comprised of Local Government Authority (LGA) Water, Sanitation and Hygiene (WASH) unit staff and sometimes WaterAid Nigeria staff – go to the community at the agreed location. If the team sees that not enough people turn up at the set time, they try to gather more people, with the support of community leaders, going to people's houses or busy areas. Between 45 minutes and one hour is spent trying to gather more people. If attempts to gather people fail (i.e. the team agrees that an insufficient proportion of the community are present) after one hour, then the triggering is cancelled. The team apologises to the people who have turned up, and requests that they mobilise more people in the future. This means that at least four people spend at least four hours (two hours in the community and two hours travelling to the community) in this phase.
4. *Follow up.* Regular community monitoring visits assess progress towards open defecation free (ODF) status, movement up the sanitation ladder and use of facilities. It is suggested that this happens on a weekly basis for each community, and each visit is estimated to take around two hours.

All 182 (97 in Ekiti and 85 in Enugu) study localities (see Section 2.3 for more details on the study design) were approached by the local implementing partners initiating the planning and mobilisation stages just described. However, the triggering stage was not achieved in all cases. Implementing partners faced obstacles in achieving the necessary level of community mobilisation in 18 out of 79 study communities in Ekiti state. On the contrary, all 85 study communities in Enugu were successfully triggered.

In the study communities where CLTS triggering failed, the team was always able to talk to and engage the community leader as part of the second step. However, the third step, which is reliant on the mobilisation of a sufficient number of community members, failed in these 18 communities, preventing the delivery all subsequent activities. We can generally identify several reasons why communities might be difficult to mobilise for CLTS activities: for example, limited attendance at the triggering due to a busy harvest period, a community leader with little mobilisation power, and seasonal migration. However, in the context of these 18 study communities, the reasons stated by programme staff were all phrased around the community's 'more urban nature'.

The fact that not all study communities went through the full set of intervention activities has implications on the interpretation of the findings we present. We discuss these implications in Section 2.5, after we have described our main outcomes of interest and have provided more details on the methodology for analysis.

2.2 Data collection

To conduct this evaluation, detailed data on sanitation behaviour, beliefs and investment decisions were elicited from households. A panel of 4,500 households has been approached in three survey waves so far: one before and two after CLTS was implemented. Implementation took place between January and June 2015. The survey round before CLTS implementation, the baseline survey (BL), included a rich data set with information on communities and households and it was conducted in November/December 2014.¹ A year later, during November/December 2015, a first rapid assessment of households (RA1) was carried out. This served as our first post-treatment survey round, which focused mainly on outcomes, and captured sanitation uptake between six and twelve months after the CLTS intervention occurred.² The second short follow-up household survey (RA2) was conducted in April 2017, around two years after the implementation of CLTS.

2.3 Research design

The impact evaluation was designed as a cluster RCT. The unit of randomisation was defined based on the realities of the programme implementation. After a detailed census of the research area, WaterAid Nigeria defined geographical units within bigger communities and small towns in which they expected to be able to implement steps 1–3 of the CLTS approach without neighbouring community members being likely to join or hear about the activities. In order to have geographical units of similar population size and with the idea of minimising spillovers from CLTS, WaterAid Nigeria therefore first created 231 such geographical areas, which we call triggerable units (TUs) in reference to the CLTS triggering event.³ The TUs are in fact a collection of villages, each containing usually between one and two villages each, with a single occasion where a TU contains seven

¹ The baseline report of the project provides a detailed analysis of these data (Abramovsky et al. 2015).

² The results of this survey round were presented in the first Rapid Assessment report (Abramovsky et al. 2016).

³ Accurate GPS data were not available at the time of designing TUs. Using their knowledge of the regions of study, WaterAid Nigeria designed them with the intention of keeping buffer villages between each TU, to prevent spillovers. As we will see in Section 4, this was only partially successful.

villages.⁴ We then randomly assigned TUs to one of two groups in order to have two observationally identical subsamples. One of these groups was subject to the CLTS intervention and the other was not. Figures A.1 and A.2 in the Appendix show the location of treatment and control TUs in Enugu and Ekiti, respectively. As described extensively in the baseline report, the randomisation was successful, in that CLTS and control groups were found to be observationally equivalent, on average, both at the household level and at the village level. This allows us to conclude that any post-treatment difference between the groups cannot be attributed to pre-treatment differences (what development economists call ‘selection bias’), but must be due to the treatment as described in, for instance, Duflo et al. (2008). For easy reference, in Table 2.1 we present the balance checks performed before the intervention took place, between the two groups in our study sample: CLTS (treatment) and control, using the whole sample at baseline.

Table 2.1. Balance of treatment and control groups.

	Control	Treatment	<i>p</i> -value
Toilet ownership			
HH has (or is constructing) a latrine (%)	37.52	37.49	0.99
HH has a functioning latrine (%)	36.19	35.87	0.92
HH has a functioning, improved toilet (%)	32.68	33.01	0.91
Toilet usage			
All members of HH use toilet (%)	34.09	33.78	0.91
At least one member of HH performs OD (%)	61.66	61.22	0.89
Head characteristics			
HH head age	55.60	54.32	0.15
HH head male (%)	64.04	62.47	0.38
HH head employed (%)	76.79	76.04	0.69
Highest education level achieved by HH head	1.439	1.451	0.88
HH size	3.991	3.733	0.03**
Children under the age of 6	0.486	0.472	0.69
Household characteristics			
HH primary activity is farming (%)	45.05	48.69	0.32
HH income, past year (thousand US\$)	0.528	0.574	0.25
HH has any savings (%)	22.50	22.73	0.92
HH has some kind of debt (%)	20.63	19.50	0.50
Owner of HH (%)	62.08	64.04	0.56
Rented (%)	15.10	14.00	0.63
Observations	4,667		

Note: Highest education level achieved by household (HH) head has five possible values: 0 = none; 1 = primary; 2 = junior secondary; 3 = senior secondary; 4 = tertiary or above. Errors are clustered at the level of randomisation (TU). Stars indicate statistical significance: * 10%, ** 5% and *** 1%. All variables measured at baseline, from both household- and village-level surveys. This table is reproduced from Abramovsky et al. (2015).

⁴ A detailed description of TUs is available in Section 2.2 of Abramovsky et al. (2015).

Table 2.1 shows the baseline means for our main outcomes of interest and a wide array of household characteristics. Importantly, it compares the means of these values between treatment and control groups, in order to verify that the random allocation of treatment among TUs achieved two identical samples, on average. We carry out individual t -tests for each variable, as well as a joint F -test for significance of the whole set of variables in predicting treatment. Overall, the results suggest that the three samples are highly balanced. Outcomes, in both ownership and usage terms, are highly balanced, and we find a single statistically significant difference between the two samples in terms of household characteristics: household size. When we conduct a test of joint significance for all these baseline characteristics in predicting treatment status, we cannot reject the hypothesis that they are insignificant (and all their coefficients are equal to zero). The F -statistics (p -values) are 1.44 (0.1123). This is consistent with very few individual statistically significant differences in certain characteristics when examined separately. Indeed, if we remove household size for this set of variables and again run a test of joint significance for predicting treatment status, we obtain an F -statistic (p -value) of 0.84 (0.64).⁵

This evidence, together with the more extensive balance checks performed in a preceding baseline report (Abramovsky et al. 2015), suggest that CLTS and control samples are comparable, as long as the only observable difference, household size, is accounted for in the analysis.

2.4 Empirical strategy

Identifying the causal impact of an intervention on a set of outcomes is typically a very challenging exercise. However, given the study design and the commendable cooperation and adherence to this design by WaterAid Nigeria and their implementing partners, we are able to attribute changes in sanitation uptake to the intervention. The design is described in detail in Abramovsky et al. (2015). Here, we provide just a very brief summary.

In our first specification, we do not distinguish between short- and medium-run impacts and we pool the observations from both follow-up waves. We compare average outcomes between CLTS and control households after the treatment, controlling for their outcomes at baseline, as follows:

$$(1) \quad y_{i,v,g,t} = \gamma CLTS_v + \theta y_{i,v,g,0} + X'_i \beta + \delta_t + \mu_g + \epsilon_{i,v,g,t}.$$

Here, $y_{i,v,g,t}$ is the outcome variable for household i , from TU v , located in LGA g , measured at RA $t = \{1, 2\}$. $CLTS_v$ is an indicator variable equal to 1 if cluster v was assigned to CLTS, and 0 otherwise. The coefficient of interest is γ , which denotes the causal impact of the CLTS treatment. $y_{i,v,g,0}$ is the value of the outcome variable measured at baseline, and its inclusion allows us to estimate causal effects more precisely than using a difference-in-difference (DID) specification (McKenzie 2012). X'_i is a vector of household characteristics, which include household size, the only unbalanced observable characteristic, as seen in Table 2.1. Finally, we allow for a time trend δ_t and for LGA fixed effects μ_g to remove level differences across LGAs.

⁵ Not shown, but available on request. Balance between CLTS and control samples when considering village-level characteristics is available in Section 2.3.1. in Abramovsky et al. (2015).

An alternative approach to ours is the one used by Cameron et al. (2015), who use DID specifications but run it only on the subsample of households that did not have a toilet at baseline. We believe that our approach is a more comprehensive approach because, besides persuading non-owners to construct toilets, CLTS informs households who already own a toilet about the importance of its maintenance and usage. At baseline, more than 70% of the toilets in our sample were pit latrines of different sorts. These pits require regular emptying (annual or biannual, generally), and will sometimes collapse and become unusable. It is therefore a margin that we think should be contemplated in our estimations, which is why we include the whole sample of households and control for baseline outcomes.

Specification 1 makes the assumption that the impact of the intervention on our outcomes of interest was identical on both post-treatment survey waves. One could pose, however, that impacts could change over time. For example, households persuaded by CLTS to build toilets between baseline and RA1 could subsequently persuade even more households within their village or network to comply and build more toilets. In this example, CLTS would have a snowball effect over time, and its effects would be larger the longer the time elapsed since the intervention. The opposite scenario is also possible: CLTS impacts observed at RA1 could vanish over time, leaving little trace after two years. This could happen if CLTS operates in the short term increasing the level of sanitation uptake but not affecting long-term trends. With time, the households in control will catch up and the effect will become smaller. This would have important implications regarding the interpretation of the impacts observed. Therefore, in a second specification, we allow impacts to vary by post-treatment survey wave:

$$(2) \ y_{i,v,g,t} = \sum_{t=1}^2 \gamma_t (CLTS_v \times I_t) + \theta y_{i,v,g,0} + X_i' \beta + \delta_t + \mu_g + \epsilon_{i,v,g,t}.$$

Now, each observation has a time subscript ($t = \{1, 2\}$) according to the survey wave they belong to: RA1 is $t = 1$ and RA2 is $t = 2$. Because we have two post-treatment survey waves (RA1 and RA2), we now have two coefficients of interest, γ_1 and γ_2 , which represent CLTS impacts at RA1 and RA2, respectively. We include survey round fixed effects δ_t to control for change in toilet ownership coverage or in other outcomes of interest – not attributed to the intervention – over time in both treatment and control areas. With this second specification, we will be able to verify whether any impacts found in the first post-treatment survey round persist in the second, using a Wald test of equality of both coefficients.

2.5 Triggerability of villages and interpretation of coefficients

Before presenting and discussing our findings, there is one more point to make, which relates to the fact that not all units selected for treatment were fully treated. We described in Abramovsky et al. (2016) how a set of communities, primarily in Ekiti, did not go through all steps of the CLTS intervention. Especially when the third step, the actual triggering activities, is not conducted, this is typically understood as CLTS not having taken place.

As we do not know which communities in our control group would not have reached the triggering stage, we cannot include in our sample only the triggered communities. Rather, we conduct our analysis with all communities initially selected to be part of the intervention, irrespective of whether they were eventually triggered or not.

In other word, our impact estimates (γ) will not actually measure the impact of CLTS itself but what is defined as the intention-to-treat (ITT) effect.⁶ This is a parameter of interest in itself, given that it measures the actual effect we can expect if we were to scale up CLTS to other villages, LGAs or states. It is reasonable to assume – and in accordance with a long history of anecdotal evidence – that full compliance will never be achieved and therefore that the ITT is a better measure of the expected benefit of the programme than an actual measure of CLTS impact on any given TU.⁷

⁶ See, for instance, Duflo et al. (2008) for a detailed technical discussion.

⁷ In Abramovsky et al. (2016) we also run a regression where we define treatment as successful triggering and we instrument this variable with the randomised treatment. Results are very comparable.

3. Impact Evaluation

3.1 Attrition

In every panel survey, where households are repeatedly interviewed over time, researchers face the problem of losing some respondents along the way. There are two main reasons why this could happen in our baseline sample of households during the short follow-up household surveys, RA1 and RA2. The first possibility is that the household cannot be located or the community itself cannot be reached by the interviewers. As mentioned above, this could be because the household moved or for reasons that might make survey work difficult or even dangerous. For example, during both RA1 and RA2, civil unrest in Ekiti South West LGA prevented interviewers from accessing one of the communities that was under government curfew. A second possibility is that the household was located, but the respondent refused to accept the interview, or provided too little time to complete the questionnaire. For our purposes, interviews that could not be carried out for any reason, whether refusal or non-availability, will be considered attrition and will not be part of our impact estimation.

Overall, our study showed very low levels of attrition. At RA1, only 148 households out of our sample of 4,671 households surveyed at baseline failed to complete interviews (3.17%), and this was similar in both CLTS (3.76%) and control groups (2.57%). Again, this includes observations lost through any of the possible reasons mentioned above. At RA2, two years and four months after the baseline survey was fielded, attrition rose to 453 households (9.70%), but this was also seemingly balanced between CLTS (10.39%) and control (9.01%) groups. The fact that we retained more than 90% of the households in our initial sample is a reassuring sign.

While, in general, attrition reduces the power a study has to identify causal effects of a policy, a larger threat is that the composition of treatment and control groups might change if attrition is asymmetric between these groups. A simple example illustrates this potential problem. Suppose that, after baseline, a large number of households with no toilets are displaced from one of our CLTS areas, and cannot be located during the RA1 survey. At the same time, suppose that all households in control areas were successfully located and interviewed at RA1. Even if CLTS had no effect at all on toilet construction, we would find that at RA1, (interviewed) CLTS households are, on average, more likely to own a toilet than control households. In this hypothetical case, the finding would be driven by changes in the sample due to attrition, not by the genuine impact of the programme on CLTS households.

Therefore, before proceeding with our impact analysis, it is important to check whether our sample is still balanced, and if attrition is correlated with treatment assignment. To do this, we check whether treatment status has any predictive power in explaining sample attrition, conditional on baseline characteristics. If this is so, then it would suggest that attrition was, in fact, asymmetric when conditioned on observable characteristics, and we would have to account for this in our impact analysis. Table 3.1 presents the results of these regressions.

Table 3.1. Regression results for RA2 attrition on treatment status

	Attrition in RA2			
	(1)	(2)	(3)	(4)
CLTS	0.01 (0.02)	0.01 (0.01)	0.00 (0.01)	0.01 (0.01)
LGA fixed effects	No	Yes	Yes	Yes
Household controls	No	No	Yes	Yes
Village controls	No	No	No	Yes
Attrition rate (control)	0.09	0.09	0.09	0.09
<i>p</i> -value for <i>F</i> -test on covariates			0.00	0.00
Number of observations	4,671	4,671	4,555	4,190

Note: Estimation results from regressions with attrition in RA2 as the dependent variable, and treatment status as the main coefficient of interest. Restricted to observations in RA2 only. Household controls include: gender, age, age squared, employment status and level of education of household head, household size, and farming as main economic activity. Village controls include: paved inner roads, presence of primary school, presence of hospital, village population (2014), and settlement population (2014). Errors are clustered at the level of randomisation (TU). Stars indicate statistical significance: * 10%, ** 5% and *** 1%.

The estimated coefficients for treatment status are small and not statistically significant in any of the three specifications, meaning that attrition was balanced across treatment groups. This is robust to the inclusion of LGA fixed effects, household- and village-level controls, as seen from Columns 2, 3 and 4, respectively. The row '*p*-value for *F*-test on covariates' shows the results of a test of joint significance for household-level controls, in Column 3, and for household- and village-level controls in Column 4. The low *p*-value signals that observable characteristics have significant explanatory power in predicting attrition. This is driven mainly by the age of the household head, given that households with older heads are more likely to drop from our sample (not shown). Importantly, this is not correlated with treatment status, meaning that attrition will not threaten our identification of causal impacts.

3.2 Impact of the CLTS intervention

In this section, we discuss findings on the impacts achieved on sanitation uptake and usage from the CLTS intervention. We first provide some visual evidence showing the upwards trend in toilet ownership, and we follow with a more formal regression analysis.

Visual evidence

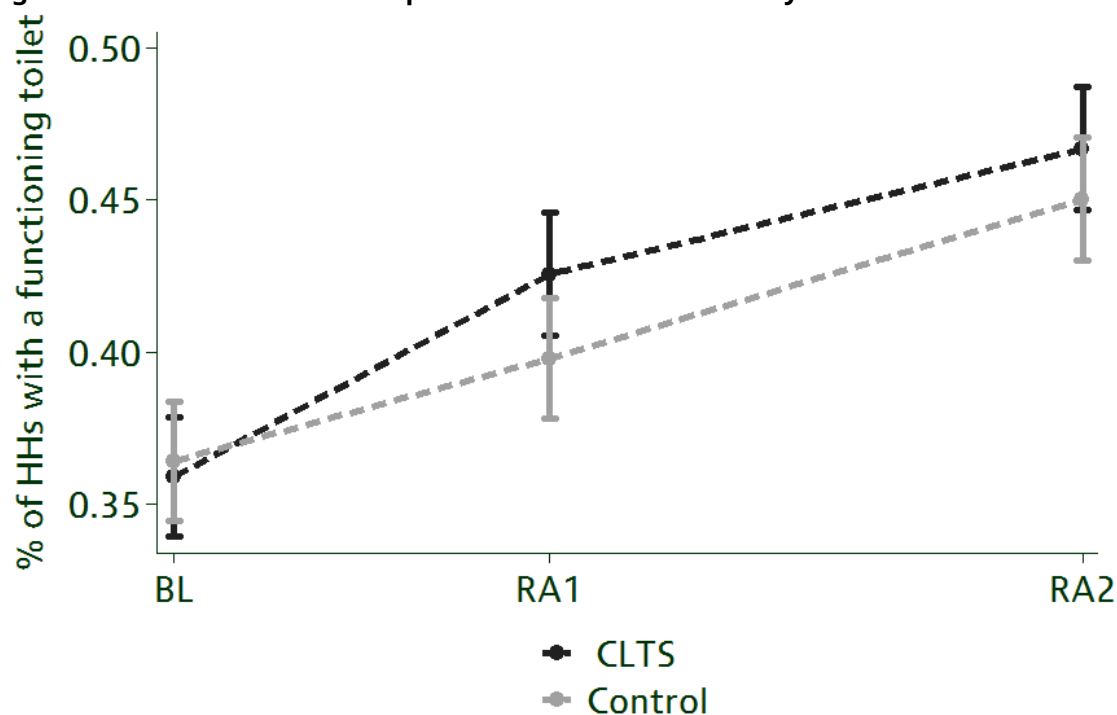
We present the average values of our main outcome of interest, namely household ownership of a functioning toilet or latrine, over the three successive survey rounds. These are unconditional means, so they might not coincide exactly with our regression analysis further below. However, they serve to take a first glance of the overall trends in our sample and our impacts.

From Figure 3.1, we see that average ownership has increased over time, for both CLTS and control groups. However, a sharper increase in ownership can be seen for the CLTS group, between baseline and RA1, which persists, albeit slightly attenuated, in RA2. Below, we perform a regression analysis to improve the statistical precision of the estimates shown in this pictorial representation of the causal impacts of CLTS and we examine

heterogeneous impacts. We also explore the impact of CLTS on other outcomes of interest.

The fact that toilet ownership increases rapidly over a period of around two years in both treatment and control areas is remarkable and somehow at odds with the long-term trends observed in Nigeria that show little progress, if any, on toilet ownership by households. We discuss these national trends in more detail, and the possible explanations for this strong increase in ownership in control areas, after the regression results, below.

Figure 3.1. Household ownership or construction of toilets by data collection wave.



Note: Lines plot the share of households owning functioning toilets of any type in each data collection period. The scale of the x-axis corresponds to the amount of time elapsed between each survey wave. The CLTS intervention was carried out shortly after the baseline survey, and at least six months before the first rapid assessment.

Overall impact estimates

Table 3.2 shows the results for the outcomes regarding toilet ownership. It shows CLTS impact estimates for ownership or construction of a toilet, ownership of a functioning toilet and ownership of an improved toilet. For each of these outcomes, we present two sets of estimates, corresponding to specifications 1 and 2 discussed above, where we estimate pooled impacts and period-specific impacts, respectively.

Let us consider in detail what we can see from the table with respect to the impacts of the CLTS intervention. In the first row, we present CLTS impacts measured by the treated coefficient γ from specification 1, which compares baseline outcomes with post-treatment outcomes, making no difference between RA1 and RA2 survey rounds. Using this method, we observe similar point estimates of around 3pp, significant at the 10% level, for all three outcomes: ownership or construction of a toilet, ownership of a functioning toilet and

ownership of an improved toilet. From baseline levels of 38%, 36% and 33%, respectively, this represents increases in toilet ownership of 8–9%.⁸ The fact that the observed impact is only marginally significant is not surprising given that study was powered to detect a change in toilet ownership variables of 8–9pp.

Table 3.2. CLTS impact on toilet ownership and construction.

Dependent var.: Toilet/OD	Constr./finished		Functioning		Improved	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated (γ)	0.03* (0.02)		0.03* (0.02)		0.03* (0.02)	
Treated $\times$ RA1 (γ_1)		0.04* (0.02)		0.03* (0.02)		0.02 (0.02)
Treated $\times$ RA2 (γ_2)		0.02 (0.02)		0.03 (0.02)		0.03 (0.02)
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
LGA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Survey round fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (BL)	0.38	0.38	0.36	0.36	0.33	0.33
F-test $\gamma_1 = \gamma_2$ (p-value)		0.29		0.79		0.70
Number of TUs	247	247	247	247	247	247
Number of HHs	4,555	4,555	4,555	4,555	4,555	4,555
Number of observations	9,110	9,110	9,110	9,110	9,110	9,110

Note: Controls include age, age squared, gender, employment status and education attainment of household head, as well as a dummy variable indicating farming as the household's main economic activity. All controls are measured at baseline. Errors in parentheses are clustered at the unit of randomisation (TU). Stars indicate statistical significance: * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$.

The second and third rows of Table 3.2 present the findings of our second specification, where we distinguish between impacts measured at RA1 (γ_1) and RA2 (γ_2). Interestingly, we find that ownership or construction of a toilet (Columns 1 and 2) increased by 4pp due to CLTS, significant at the 5% level, between baseline and RA1, but when comparing baseline and RA2, this effect falls to 2pp and is not statistically significant. A similar pattern is observed in Columns 3 and 4 corresponding to ownership of functioning toilets, where statistically significant impacts of 3pp appear at RA1 and become statistically insignificant at RA2. No significant impacts are observed for the ownership of a functioning, improved toilet using this second specification (Columns 5 and 6).

The explanation to these findings is in Figure 3.1. Average ownership in CLTS and control groups diverged significantly between baseline and RA1, showing a marked increase in CLTS villages. Between RA1 and RA2, control villages began to catch up with CLTS villages, mainly due to a dip in the rate of toilet construction among CLTS villages. This translates into no additional gains from CLTS relative to control areas by RA2, even though both exhibit an increase in toilet construction over time. These estimates show that the increases in toilet coverage attributable by CLTS during the first six to twelve months after the intervention are no longer detectable by RA2, when CLTS and control areas exhibit

⁸ Baseline averages (control mean) are shown in the lower panel of Table 3.2.

much more similar levels of coverage.⁹ The noteworthy increase in toilet coverage in control areas during this period raises the question of whether CLTS might be affecting these areas via spillovers. This is important because if spillovers were indeed playing a role here, then our estimates from Table 3.2 might be a lower bound to actual programme impacts. This will be discussed more extensively in Section 4.

Table 3.3 presents the results of a similar analysis on outcomes of toilet usage and OD behaviour. During our baseline survey, we observed that conditional on having a toilet, respondents in our sample overwhelmingly declared that they did use it. So, it is not surprising to see that CLTS did not change that, as we see in Columns 1 and 2. However, Columns 3–6 show that the increase in toilet construction detected in Table 3.2 effectively translated into an overall reduction of OD levels. First of all, looking at OD behaviour by the main respondent, in Columns 3 and 4, we see that the reduction of 4pp in OD shown in the first row of Column 3 goes hand in hand with the 3pp increase in ownership of functioning toilets. This effect appears to be persistent over time, as we can observe statistically significant reductions of 5pp and 4pp at RA1 and RA2, respectively, in Column 4. Columns 5 and 6 show the results for an outcome only measured at baseline and RA2: OD practised by any member of the household. A statistically significant reduction in OD practice of 4pp is seen, between baseline and RA2. In other words, it appears that the modest increases in toilet construction and ownership presented in Table 3.2 reduced OD levels in our area of study by a comparable, and sustainable, amount.

Table 3.3. CLTS impact on toilet usage and OD behaviour.

Dependent var.: Toilet/OD	All use		OD (main resp.)		OD (any member)	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated (γ)	0.03 (0.02)		-0.04** (0.02)		-0.04* (0.02)	
Treated $\times$ RA1 (γ_1)		0.02 (0.02)		-0.05** (0.02)		0.00 (-)
Treated $\times$ RA2 (γ_2)		0.03 (0.02)		-0.04* (0.02)		-0.04* (0.02)
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
LGA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Survey round fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (BL)	0.34	0.34	0.61	0.61	-	-
F-test $\gamma_1 = \gamma_2$ (p-value)		0.78		0.54		0.06
Number of TUs	247	247	247	247	247	247
Number of HHs	4,555	4,555	4,542	4,542	2,278	2,278
Number of observations	9,110	9,110	9,084	9,084	4,555	4,555

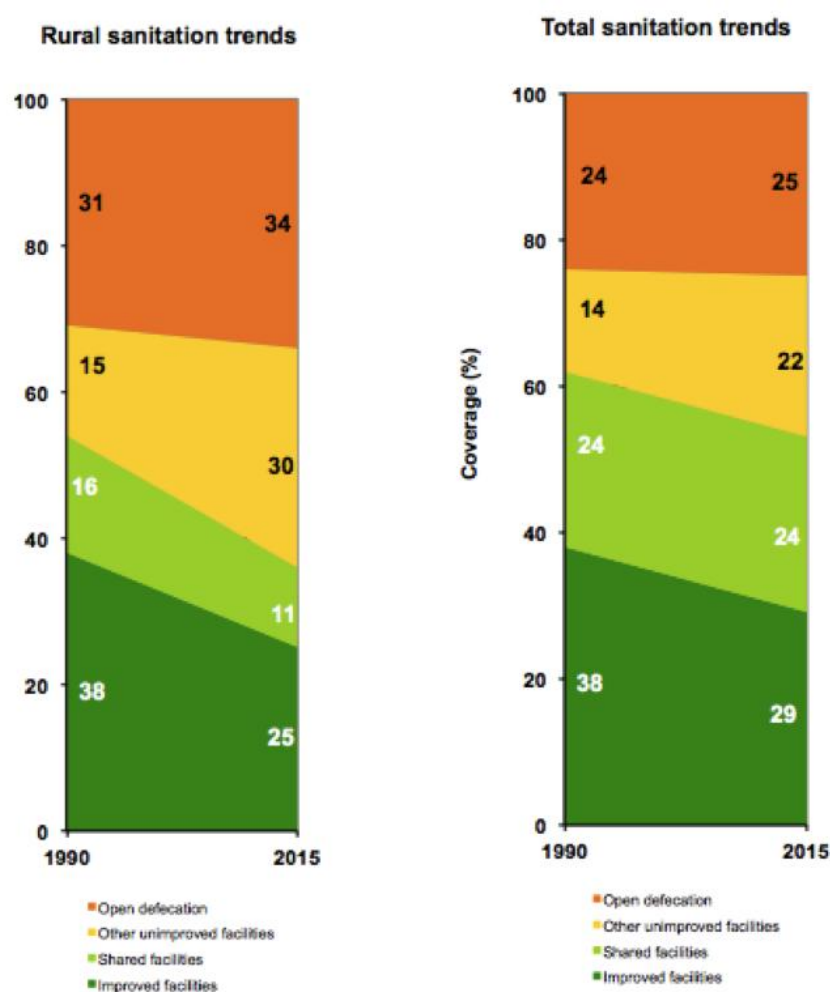
Note: Controls include age, age squared, gender, employment status and education attainment of household head, as well as a dummy variable indicating farming as the household's main economic activity. All controls are measured at baseline. Errors in parentheses are clustered at the unit of randomisation (TU). Stars indicate statistical significance: * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$.

It is worth putting the findings on uptake and usage into the wider context of Nigeria. The right panel of Figure 3.2 shows that over the 25-year-period between 1990 and 2015, OD

⁹ In the row 'F-test $\gamma_1 = \gamma_2$ (p-value)', we test the hypothesis that both coefficients are simultaneously equal to zero, and we cannot reject this to a 10% degree of confidence. This does not necessarily imply that they indeed are the same, but it points towards caution when interpreting this structure of impacts over time.

stayed constant in the country as a whole. In 1990, 24% of households practised OD, implying that 76% used some type of toilet; in 2015, the same percentage is 25%, suggesting a minor decline in sanitation coverage in percentage terms. Of course, this translates into a significantly larger number of households with no access to sanitation by 2015, given that, according to the World Bank, the population in Nigeria almost doubled over this period. In rural areas (left panel), this decline was more pronounced, as these areas experienced an increase from 31% to 34% in OD rates over the same period. Viewed in this larger context, the positive trends in both areas (CLTS and control) are remarkable. Furthermore, the causal impact of CLTS on sanitation uptake found can be considered an important achievement.

Figure 3.2. Sanitation trends: Nigeria 1990–2015.



Source: Joint Monitoring Programme (2015).

3.3 Heterogeneous impacts

The results we have presented above summarise average CLTS impacts in our study areas. In this section, we delve into whether certain groups of people were more or less affected by the CLTS intervention. In our previous report, we documented how, by RA1, CLTS appeared to have a stronger impact on households whose heads either were not highly educated (incomplete primary school or illiterate) or were women. At the same time,

higher impacts were detected among poorer households (those with below median wealth levels at baseline), households with children below the age of 5, households with seniors (over 65 years old) and households with debts or no savings. We interpreted this as evidence that CLTS had been most effective in mobilising toilet construction and usage among vulnerable households.

In this section, we carry out a similar analysis, and we test whether these impacts in selected populations are still present by RA2. For simplicity, we focus the analysis on a single outcome: ownership of a functioning toilet. Recall that over the whole population, as we have seen in Column 4 of Table 3.2, CLTS increased the share of households owning functioning toilets by 3pp (significant at a 10% level) when comparing baseline with RA1. However, no significant effects are observed when comparing baseline with RA2.

Table 3.4. CLTS impact by characteristics of the household head.

Dependent var.: Functioning	Primary education		Male		Children		Seniors	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Yes	No	Yes	No	Yes	No	Yes	No
Treated × RA1 (γ_1)	0.02 (0.02)	0.07** (0.03)	0.02 (0.02)	0.06** (0.02)	0.04* (0.03)	0.03 (0.02)	0.06** (0.03)	0.02 (0.02)
Treated × RA2 (γ_2)	0.02 (0.02)	0.05 (0.03)	0.02 (0.02)	0.04 (0.03)	0.03 (0.03)	0.03 (0.02)	0.07** (0.03)	0.01 (0.02)
Control mean (BL)	0.40	0.28	0.38	0.33	0.37	0.36	0.38	0.35
F-test $\gamma_1 = \gamma_2$ (p-value)	0.55	0.04	0.46	0.06	0.18	0.21	0.01	0.65
Number of TUs	244	240	246	238	238	247	237	246
Number of HHs	3,104	1,451	2,888	1,667	1,354	3,201	1,460	3,095
Number of obs.	6,208	2,902	5,776	3,334	2,708	6,402	2,920	6,190

Note: Controls include age, age squared, gender, employment status and education attainment of household head, as well as a dummy variable indicating farming as the household's main economic activity. All controls are measured at baseline. Errors in parentheses are clustered at the unit of randomisation (TU). Stars indicate statistical significance: * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$.

Table 3.4 shows the results of running regressions using the specification in equation 2 on a series of selected portions of our sample of households. For example, Column 1 shows the impact of CLTS on households whose head had completed primary education at baseline, while Column 2 shows the results for households whose head had not completed it. As in Abramovsky et al. (2016), we see that positive and significant programme impacts are only observed among the second group, in Column 2, where CLTS households were, on average, 7pp more likely than control households to own a functioning toilet by RA1. By RA2, this difference was smaller (5pp, as seen in the second row of the table) and no longer significantly different from zero.

This pattern of findings is repeated over the rest of the table. In Columns 3 and 4, we detect higher impacts at RA1 among households with female heads, but these cease to be significant by RA2. Columns 5–7 show that CLTS impacts were larger among households with children (Columns 5 and 6) and households with seniors (Columns 7 and 8). In Column 7, we observe that CLTS households with seniors at baseline manifest statistically significant higher levels of functioning toilet coverage than controls at RA2, something

which is only observed for this category. Therefore, we confirm the findings in our previous report, but we observe that these impacts have faded out in most categories, over the course of the 16 months that elapsed between RA1 and RA2.

Shifting our attention to economic variables at the household level, Table 3.5 performs a similar exercise as above splitting households according to whether they report to have any savings, any debts and to their level of wealth. CLTS impacts did not vary according to whether households had any savings, but were higher among those with no debts, between baseline and RA1. These do not persist by the time of RA2. The same pattern can be seen in Column 5 for the poorer households in our sample. Here, wealth is measured by a composite index constructed using the answers to a series of questions on ownership of certain durable assets, at baseline. The poorest half of our sample at baseline shows CLTS impacts of 4pp between baseline and RA1 that eventually fade and become undetectable by RA2.

Table 3.5. CLTS impact by economic characteristics of the household.

Dependent var.: Functioning	Has savings		In debt		Below median wealth	
	(1)	(2)	(3)	(4)	(5)	(6)
	Yes	No	Yes	No	Yes	No
Treated $\times$ RA1 (γ_1)	0.03 (0.03)	0.03 (0.02)	0.00 (0.03)	0.04** (0.02)	0.04* (0.02)	0.02 (0.02)
Treated $\times$ RA2 (γ_2)	0.02 (0.03)	0.03 (0.02)	-0.01 (0.03)	0.03 (0.02)	0.03 (0.03)	0.03 (0.02)
Control mean (BL)	0.50	0.32	0.32	0.37	0.26	0.48
F-test $\gamma_1 = \gamma_2$ (p-value)	0.52	0.20	0.98	0.09	0.19	0.43
Number of TUs	222	247	222	247	247	232
Number of HHs	1,065	3,490	912	3,643	2,518	2,037
Number of observations	2,130	6,980	1,824	7,286	5,036	4,074

Note: Controls include age, age squared, gender, employment status and education attainment of household head, as well as a dummy variable indicating farming as the household's main economic activity. All controls are measured at baseline. Errors in parentheses are clustered at the unit of randomisation (TU). Stars indicate statistical significance: * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$.

This set of findings suggests that CLTS is most effective among a specific group of households that could be thought of as the most vulnerable. These households are vulnerable in several dimensions: their household heads have no education at all (primary school not completed), they have children and senior members (making them more exposed to disease) and they are relatively less wealthy than the rest of the households in the sample. Consequently, they are also less likely to have access to financial instruments (no debts). The reason why CLTS affects this population in particular goes beyond the scope of this report, and will be explored in greater detail in a separate paper, dedicated entirely to understanding the channels of impact of this intervention.

An important finding here is that these impacts, as well as those for the general population shown in the previous section, are no longer observable by RA2, while the uptake of toilets increases in both CLTS and control areas. This motivates our analysis into the possibility of programme spillovers from CLTS to control areas, in Section 4.

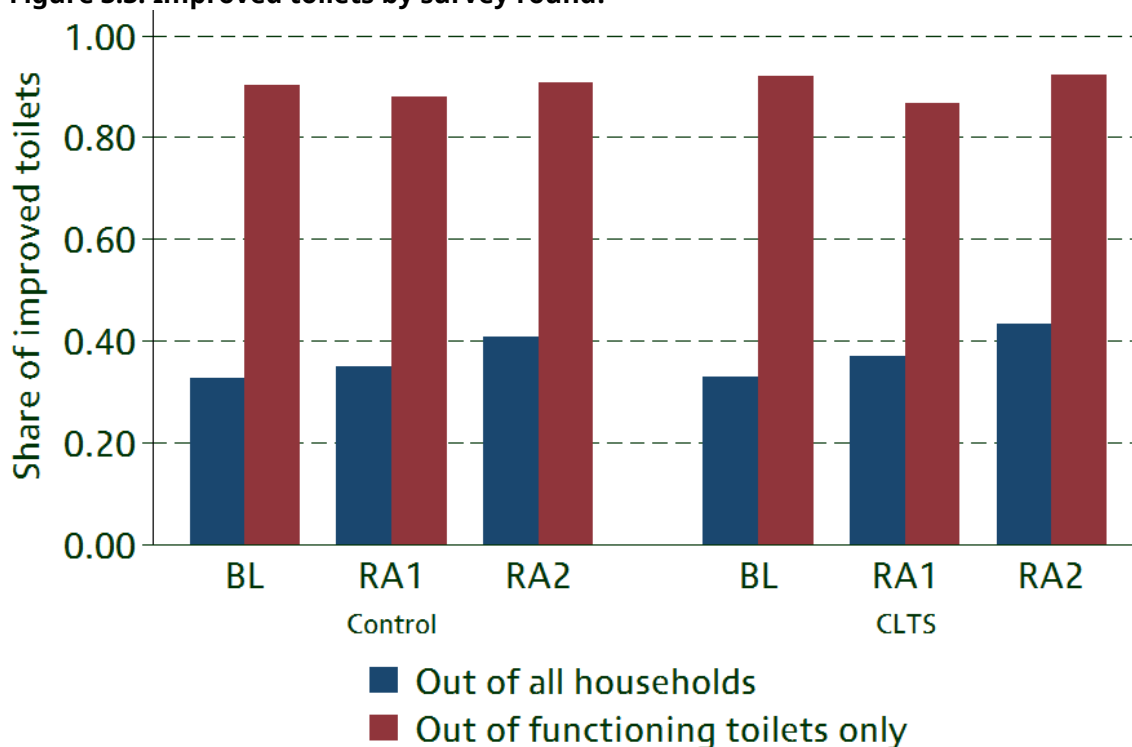
3.4 Improved sanitation

Before moving on, however, a reasonable concern raised in our first Rapid Assessment report focused on the quality of the additional toilets constructed in CLTS areas between baseline and RA1. While CLTS had a positive and significant effect on toilet construction and ownership, this effect was not as large for improved toilets specifically, as seen in Columns 4 and 6 of Table 3.2. This suggests that the share of unimproved toilets might have increased slightly in CLTS areas, albeit not in a statistically significant way.

This could be a problem for at least two reasons. First, while toilet use is considered safer than the practice of OD, unimproved toilets do not provide such an effective barrier between people and faeces, and are therefore unlikely to provide the same (or, in the worst case, any) health benefits as safe toilets. At the same time, unimproved toilets tend to depreciate faster than improved toilets. If the toilets built because of CLTS had a higher proportion of unimproved units than those constructed in control areas, then CLTS areas would, on average, experience a higher rate of depreciation or, in other words, a higher share of toilets going out of use over time. We could then expect to see coverage levels between these two areas converging over time, as these additional unimproved units stop functioning and go out of use, and the improvements in toilet coverage due to CLTS estimated at RA1 are slowly undone.

Figure 3.3 plots the evolution of two variables over the course of the three successive data collection waves, by treatment group. The first bar (blue) shows the share of all households who own an improved toilet. We observe no statistically significant difference in the evolution of CLTS and control groups, something that we have already found in Columns 5 and 6 of Table 3.2. This tells us that CLTS had no impact on the share of households owning improved toilets.

Figure 3.3. Improved toilets by survey round.



Note: Blue bars show the share of households who own functioning improved toilets out of all the households in our sample. Red bars show the share of existing, functioning toilets that are considered improved. Improved and unimproved toilets are classified according to UNICEF and World Health Organization (WHO) Joint Monitoring Programme criteria (UNICEF and WHO 2015).

Source: Authors' own calculations based on baseline, RA1 and RA2 questionnaires.

The second bar (red) in Figure 3 aims to answer the concerns raised in Abramovsky et al. (2016) and described above. It plots the share of households with improved toilets out of the total number of households who own any kind of functioning toilet. If CLTS had affected the mix of improved/unimproved toilets in CLTS areas, then we would observe a different trend of the red bars between control and CLTS households. On the contrary, while there appears to be a small dip in improved toilets at RA1 among CLTS households, by RA2 these shares are identical among treatment arms. The reason for this could be that CLTS does not affect the mix of improved/unimproved toilets being built in a region or, alternatively, that this effect is small and undetectable by RA2, when CLTS impacts on construction of toilets of any kind have already faded.

These findings imply that, although CLTS seems to have induced the construction of more unimproved than improved toilets, 14 months after programme implementation, at RA2, this does not seem to have been counterproductive in terms of slippage back to OD.

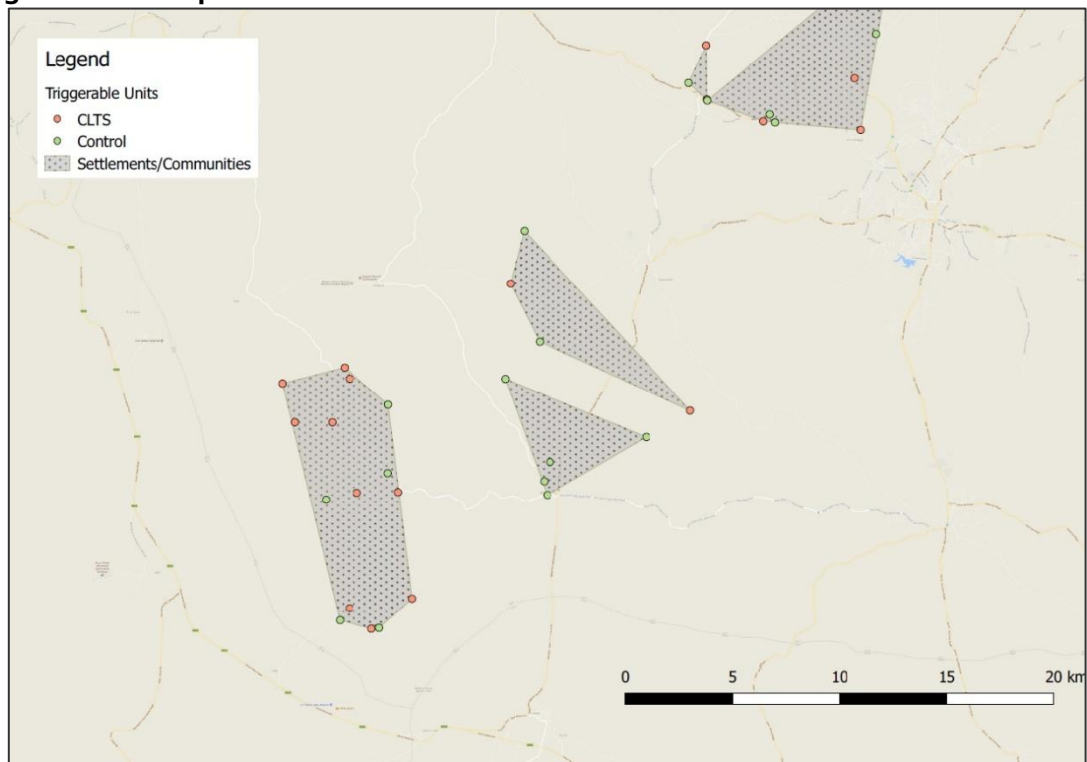
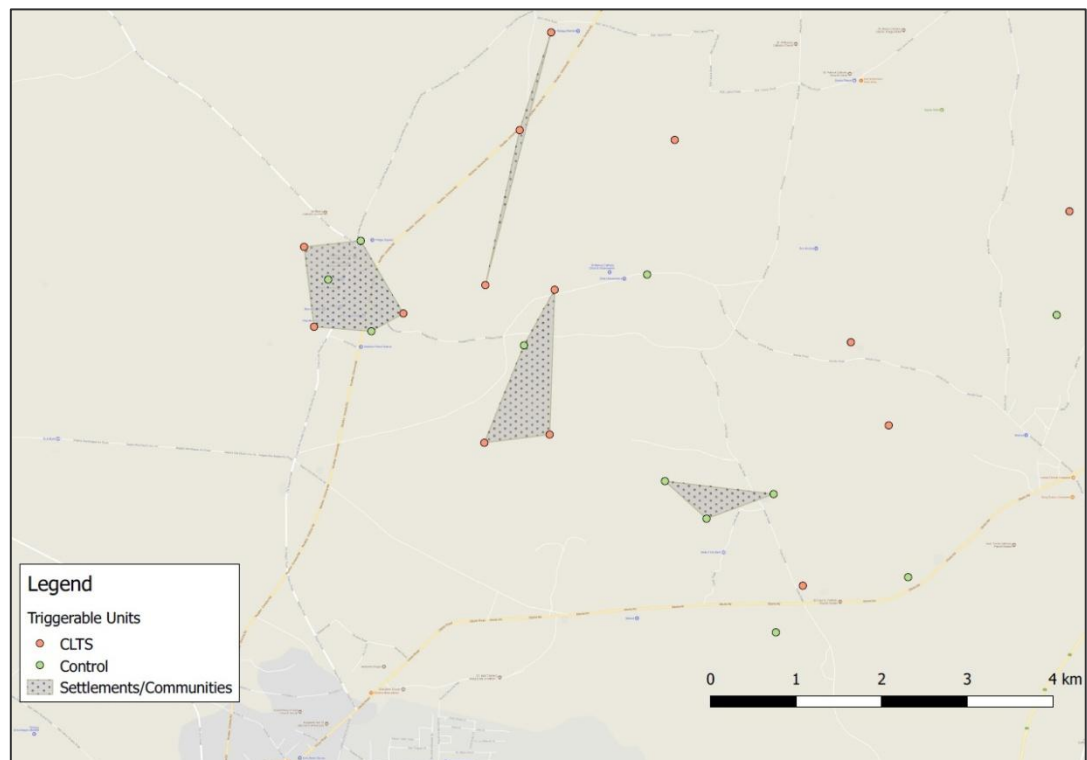
4. Programme Spillovers

The CLTS intervention included a significant information component: households were informed about the health risks involved in OD as a way of incentivising toilet construction and usage. Information travels fast and one could pose that, sooner or later, advice that is deemed useful by households might spill over to neighbours, relatives and acquaintances. While this might be positive from the point of view of a policymaker, it may bring challenges to the evaluation of the policy's impact. This is because information spillovers to control areas turn them into imperfect counterfactuals. To give an extreme example, if an information campaign has a significant effect on behaviour, but the information reaches all the individuals in the control group as well, then the researcher will not be able to find any detectable differences between treatment and control outcomes and will conclude that the treatment is ineffective.

The DID specification used so far compares average outcomes in the CLTS group with those of the control group. Ideally, the control group represents the outcome we would expect to observe from CLTS households, had the CLTS intervention not occurred. As discussed above, the random assignment of households in a sample to either treatment or control is the benchmark in ensuring that both groups are identical. However, there is always the risk that the assignment of treatment to one household affects the outcomes of other households. This is a violation of what is called the stable unit treatment-value assumption (SUTVA), and it is more common in social experiments than in drug trials. To mitigate this risk, social experiments usually treat groups of subjects (e.g. households in villages) and, where possible, choose villages that are not likely to be socially and economically linked (e.g. sometimes imposing some geographical distance as a proxy for restricting social and economic interactions). If one thinks that there could be spillovers, then average outcomes in the control group will not be representative of the true counterfactual situation, for example, the situation in which CLTS did not take place. Thus, our control areas will perform better than they would, due to these (unaccounted for) indirect impacts of CLTS, and therefore will not be identical to households who were never exposed to CLTS or its spillovers. This will bias our estimates of impacts downwards, underestimating the true effect of the programme.

Randomisation into CLTS was carried out at the unit of the TU, which is a geographical unit that does not correspond to a specific administrative unit but was designed by WaterAid, as discussed in Section 2. TUs are smaller than or equal to communities, meaning that some communities might be composed of a single or several TUs.¹⁰ This resulted in some communities having all their villages (and households) assigned to CLTS, others having just a few of them assigned, and others having no CLTS villages (and households). Figure 4.1 uses the LGAs of Ekiti South West and Igbo Eze North as examples. The figures show the centroid locations of TUs (orange for CLTS and green for controls) in our study, and show settlements as grey shaded areas. Settlements composed of a single TU are simply presented as dots. As can be seen, while some settlements contain only CLTS or control TUs, others are composed of TUs from both treatment groups.

¹⁰ Community is the term used in Enugu, whereas in Ekiti, the equivalent is settlement.

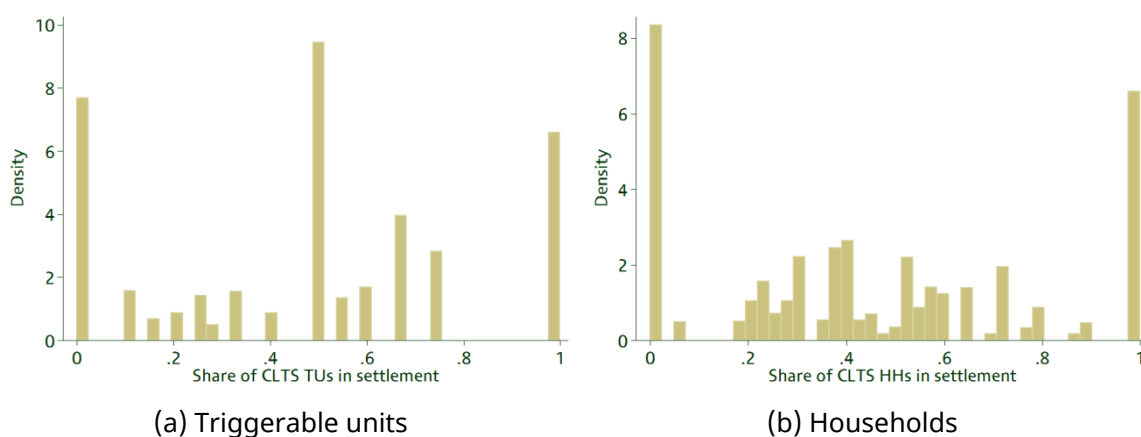
Figure 4.1. Example of settlements with different shares of CLTS TUs.**(a) Ekiti South West LGA****(b) Igbo Eze South LGA**

Note: The maps show a section of the LGAs of Ekiti South West and Igbo Eze South. Settlements are shown by shaded areas, and are composed of multiple TUs. Designed using information gathered in RA2.

So, while a significant effort was made to avoid spillovers, in some cases TUs assigned to CLTS and control fell within the limits of the same administrative unit (settlement or community). In principle, the TUs were designed in a way that would minimise these spillovers, even in cases where they fell within the same administrative units. However, given that some TUs share some institutions and/or markets, it is worth exploring whether or not there is evidence supporting the existence of spillovers.

To do this, we start by showing in Figure 4.2 the outcome of this randomisation at the TU and household level in relation to the administrative units. While there are a significant number of administrative units that contain only CLTS or Control households, 49 out of the 105 (49%) settlements in our sample contain both CLTS and Control households. Of the settlements that are 100% CLTS, 75% contain a single TU, while this is the case for 63% of the exclusively Control settlements. Apart from these extremes, the Figure shows that our randomisation created settlements with varied levels of CLTS coverage.

Figure 4.2. Share of households and TUs assigned to CLTS by settlement.



Note: Total number of households and TUs in each settlement calculated from the mapping exercise carried out during mid-2014, and including CLTS, control and areas not included in our study.

We explore two ways to estimate the magnitude of these potential spillover effects. First, we exploit the fact that, as a result of our research design, settlements or autonomous communities may be composed of CLTS, control or both types of households. Because TUs are smaller than settlements, and because treatment was assigned at the TU level, our sample consists of settlements with varying combinations of CLTS and control households. Therefore, this first approach takes a wider definition of treatment, and considers any household living in a settlement with at least one CLTS household as a treated household. As we can see from Table 4.1, this group is composed of households in TUs originally assigned to CLTS and households in TUs that were assigned to control but that are in settlements with CLTS TUs. We refer to this redefined treatment group as 'CLTS (+ spillovers)'. Our assumption here is that information flows freely within settlements but not across them. While this is, of course, not true in the long term, we believe it is a reasonable assumption to make for the extent of the study period. However, the control group in this case will be composed of households in exclusively control settlements (i.e. settlements where no households were assigned to CLTS). We refer to this group as 'pure controls'. The result of comparing outcomes between these two groups is called the total

causal effect (TCE) as defined by Baird et al. (2016), also referred to as the overall causal effect by Hudgens and Halloran (2008).¹¹

Table 4.1. Treatment status at different levels.

Level	Treatment status	
Settlement/autonomous community	CLTS (+ spillovers)	Pure controls
Triggerable unit	CLTS control	control

Note: Randomisation carried out at the TU level. Source: Adapted from Özler (2016).

This approach has some important caveats. Because the pure control group is composed mainly of small settlements containing a single TU, they might not be an identical counterfactual group. This asymmetry in size is a mechanical result of our definition of the CLTS (+ spillovers) group. Given that it only takes one CLTS TU in a community for it to be classified as CLTS (+ spillovers), the larger the community, the more unlikely it is to be a pure control. A community composed of a single TU has a 50% chance of being assigned to CLTS (and therefore be part of the CLTS (+ spillovers) group) or control (and therefore be part of pure controls). A community composed of ten TUs has a much lower chance of ending up as part of the pure control group, given that all its TUs should have been assigned to control for this to happen. As noted in Özler (2016), in order to appropriately estimate the effects we are after, we should perform a two-step randomisation whereby, first, settlements are assigned a certain treatment saturation level (such as 0, 50% or 100% CLTS) and, then, treatment can be randomly assigned within each settlement to achieve these target levels. By not doing this, we risk assigning treatment saturation levels non-randomly, thus hindering the comparability of treatment and control groups.

Indeed, when we test for balance across these two groups to see whether there are any significant observable differences between them (Table 4.2), we find that, together with a few differences at the household level, the two groups are significantly different in the dimension expected. Pure controls are likely to belong to smaller settlements, and their villages are less likely to have a hospital. Therefore, we include these covariates as additional controls when repeating our impact regressions with this new definition of control areas.

In an alternative approach, we tackle the issue of the different composition of our treatment and control samples more directly. In order to avoid the asymmetry in the composition of CLTS (+ spillovers) and pure control groups, we define a third, pure CLTS group, composed only of settlements in which all TUs were assigned to CLTS. We then compare average outcomes at RA1 and RA2 between these pure CLTS and pure control groups only, and we discard all the households belonging to settlements that contain both CLTS and control households. These two groups should, in principle, be more similar to each other, since the probability of each kind of settlement to belong to either of these groups is identical.

¹¹ Importantly, Baird et al. (2016) state that the TCE can be identified only if the treatment saturation level (i.e. the share of treated units in each cluster) is randomised first, and only then can the treatment be assigned randomly within each cluster. This was not the case in our design as we expected spillover not to be present, so our estimates must be taken with caution. In future work, we will perform multiple robustness checks to provide more precise bounds to the true treatment effects in the presence of spillovers.

Table 4.2. Balance of CLTS (+ spillovers) and control groups.

	Pure control	CLTS (+ spillovers)	p-value
Toilet ownership			
HH has (or is constructing) a latrine (%)	36.02	37.89	0.71
HH has a functioning latrine (%)	35.07	36.29	0.80
HH has a functioning, improved toilet (%)	30.21	33.47	0.49
Toilet usage			
All members of HH use toilet (%)	32.46	34.31	0.70
At least one member of HH performs OD (%)	63.86	60.86	0.55
Head characteristics			
HH head age	55.96	54.73	0.42
HH head male (%)	63.71	63.18	0.83
HH head employed (%)	76.05	76.53	0.84
Highest education level attended by HH head	1.242	1.489	0.03**
HH size	4.605	3.696	0.00***
Children under the age of 6	0.620	0.448	0.00***
Household characteristics			
HH primary activity is farming (%)	51.20	45.96	0.35
HH income, past year (th. US\$)	0.469	0.570	0.07*
HH has any savings (%)	20.07	23.17	0.30
HH has some kind of debt (%)	22.99	19.39	0.21
Owner of HH (%)	73.82	60.66	0.02**
Rented (%)	7.346	16.17	0.01***
SanMark treatment status (%)	48.93	47.55	0.87
<i>F</i> -test	$F(18, 104) = 1.46$		
<i>p</i> -value	0.12		
Village characteristics			
Has graded internal roads (%)	38.88	44.74	0.54
At least one primary school (pub/priv) (%)	61.63	69.69	0.32
Has a hospital (%)	6.641	15.86	0.10*
Village population (listing)	498.5	503.0	0.95
Settlement population (listing)	1,345.7	2,438.4	0.00***
<i>F</i> -test	$F(23, 100) = 1.69$		
<i>p</i> -value	0.04		
Number of observations	835	3,789	

Note: Errors are clustered at the settlement level. Stars indicate statistical significance: * 10%, ** 5% and *** 1%. All variables measured at baseline, from both household- and village-level surveys.

Table 4.3. Balance of treatment and control groups (pure CLTS or control settlements only).

	Pure control	Pure CLTS	<i>p</i> -value
Toilet ownership			
HH has (or is constructing) a latrine (%)	36.02	36.57	0.93
HH has a functioning latrine (%)	35.07	35.31	0.97
HH has a functioning, improved toilet (%)	30.21	31.54	0.82
Toilet usage			
All members of HH use toilet (%)	32.46	33.26	0.89
At least one member of HH performs OD (%)	63.86	63.66	0.97
Head characteristics			
HH head age	55.96	54.65	0.39
HH head male (%)	63.71	60.07	0.21
HH head employed (%)	76.05	78.24	0.40
Highest education level attended by HH head	1.242	1.341	0.42
HH size	4.605	4.055	0.01***
Children under the age of 6	0.620	0.519	0.05*
Household characteristics			
HH primary activity is farming (%)	51.20	52.99	0.78
HH income, past year (th. US\$)	0.469	0.533	0.35
HH has any savings (%)	20.07	23.36	0.33
HH has some kind of debt (%)	22.99	23.09	0.98
Owner of HH (%)	73.82	75.20	0.82
Rented (%)	7.346	8.343	0.77
SanMark treatment status (%)	48.93	45.37	0.71
<i>F</i> -test	$F(18, 63) = 2.27$		
<i>p</i> -value	0.09		
Village characteristics			
Has graded internal roads (%)	38.88	33.42	0.67
At least one primary school (pub/priv) (%)	61.63	70.34	0.41
Has a hospital (%)	6.641	8.245	0.81
Village population (listing)	498.5	482.8	0.81
Settlement population (listing)	1,345.7	1,443.7	0.68
<i>F</i> -test	$F(23, 59) = 2.91$		
<i>p</i> -value	0.00		
Number of observations	922	793	

Note: Errors are clustered at the settlement level. Stars indicate statistical significance: * 10%, ** 5% and *** 1%. All variables are measured at baseline, from both household- and village-level surveys.

Table 4.3 shows that when comparing pure CLTS and pure control communities only, these are mostly balanced. Except for an imbalance in average household size and the

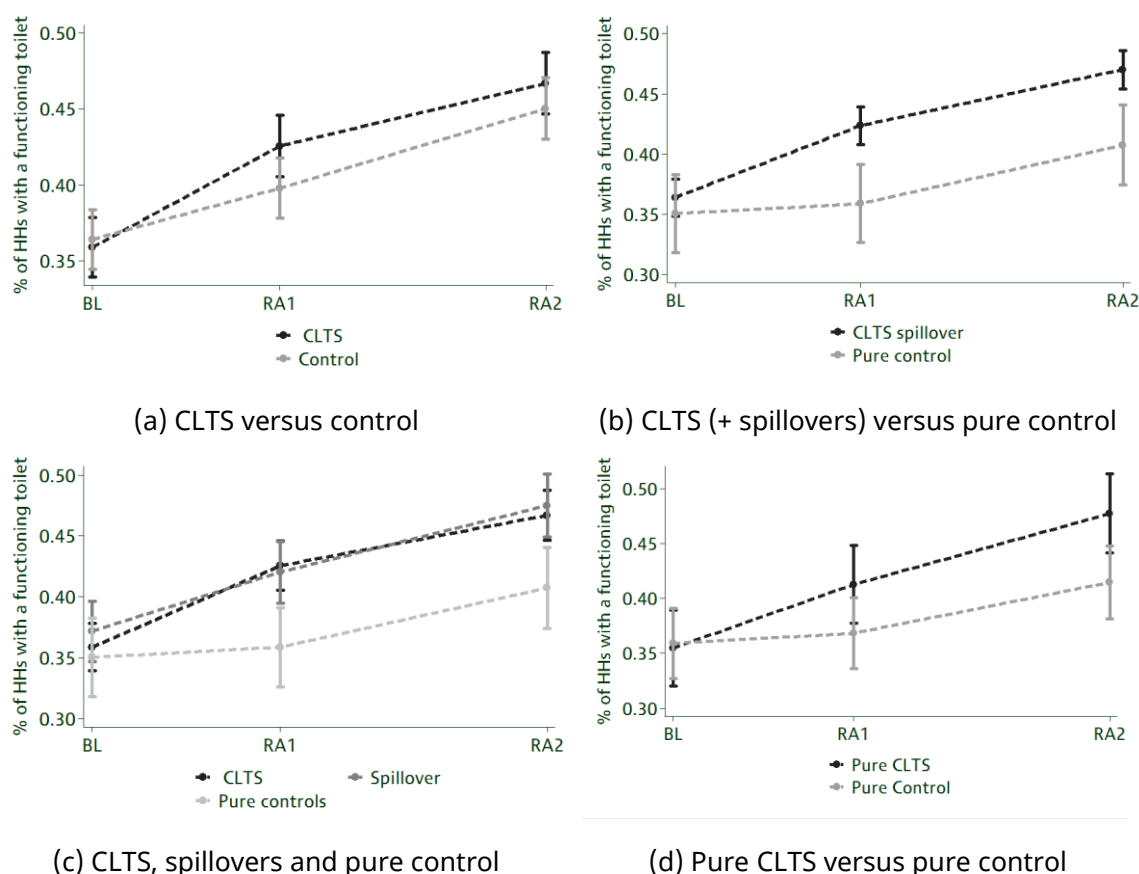
presence of children – which we control for in our impact regressions below – the two groups present no other differences. In contrast with Table 4.2, we find no observable differences at the village level.

While the disadvantage of this second approach is that it provides impact estimates only for small settlements (and hence cannot be generalised to our full sample), it nevertheless serves two purposes. First, it provides an important robustness check to the estimates from the first approach in which we compare outcomes between CLTS (+ spillovers) and pure controls. Secondly, it provides a benchmark estimation of CLTS effects abstracting from any possible spillovers, by removing any control households exposed to CLTS spillovers from the sample. By combining the findings of these two strategies, we aim to achieve a better understanding of the impacts of the CLTS programme.

4.1 Results

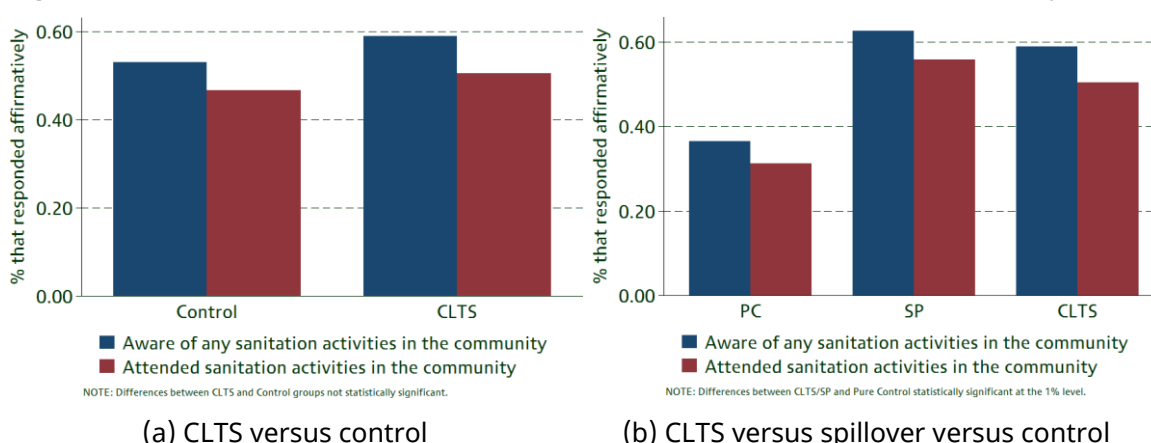
Figures 4.3(b) and (c) provide evidence consistent with the idea that control units were indeed affected by their neighbours being assigned to CLTS. In Figure 4.3(b) we use the definition of treatment described above and we see that the increase in ownership is more pronounced once we account for spillovers. Figure 4.3(c) plots the average outcomes of each of the three groups of interest separately (CLTS households, pure control households, and spillover households); that is, control households living in settlements where other households received CLTS treatment. This plot is especially revealing because it shows the almost identical path followed by CLTS and spillover groups, lending support to our alternative assumption that information flows freely (and fast) within settlements, but does not flow across settlements.

Average ownership evolved in a strikingly similar way among households assigned to CLTS and control households in settlements with CLTS households (spillovers, in Figure 4.3(c)). At the same time, pure controls evolved very differently from these two groups, with average ownership staying flat between baseline and RA1, and only starting to pick up after that. These figures suggest significant within-settlement spillovers immediately after the CLTS intervention was carried out. They show that, if present at all, across-settlement spillovers occur significantly slower. Figure 4.3(d) compares the average trends for pure CLTS and pure controls only, excluding households in partially treated settlements. The difference in the evolution of toilet coverage, particularly by RA2, is also clear, with no sign of pure controls closing the gap with pure CLTS households.

Figure 4.3. Household ownership or construction of toilets by data collection wave.

Note: Lines plot the unconditional share of households owning functioning toilets of any type. Spillovers are control households located in settlements where at least one other TU was assigned to CLTS. In (b), these households are included in the CLTS (+ spillovers) group, and the comparison group is composed exclusively of pure controls. The scale of the x-axis corresponds to the amount of time elapsed between each survey wave. CLTS was carried out shortly after the baseline survey, and at least six months before the first rapid assessment.

Figure 4.4 shows the average response at RA2 to two questions. The first (blue) bar shows the results when the respondent was asked whether they were aware of any sanitation activities carried out in the past in their communities. The second (red) bar shows the results when the respondent was asked whether they or any other member of the household attended these activities. Figure 4.4(a) suggests very similar levels of awareness and attendance between CLTS and control households. While respondents might refer to other, sanitation-related activities that are not part of our study, we should still expect higher levels of awareness and attendance in treatment areas. Figure 4.4(b) adds clarity to this: high awareness among control households is driven by those living in settlements with other CLTS households – in other words, spillover households. This group, in fact, has average awareness and attendance figures that are not significantly different from those of CLTS households. However, pure control households (left two columns in Figure 4.4(b)) have lower awareness and attendance levels, and these are significantly different from those of spillover and CLTS groups.

Figure 4.4. Awareness and attendance of sanitation activities in the community.

Note: PC denotes pure control households (i.e. control households in settlements with no CLTS households). SP denotes spillovers (i.e. control households in settlements with at least one treated household). CLTS denotes CLTS households. The share of respondents who declare that they have heard of (or attended) activities around sanitation carried out in the community.

Our findings from Figures 3.1 and 4.4 lead us to conclude that there is significant evidence for the presence of spillovers. In Table 4.4, we compare three toilet coverage outcomes for the CLTS (+ spillovers) group with those of pure controls, over the two post-treatment survey waves, using regression analysis. At first glance, we can see that pooled estimates (in Columns 1, 3 and 5) show positive and significant CLTS impacts, which are higher than those estimated in our original specification (presented in Table 3.2). Column 2 presents our estimated programme impacts on the construction or ownership of a toilet, and we see that, by RA1, these impacts are estimated to be of 6pp and significant at the 5% level. When comparing CLTS and control households in Table 3.2, this same coefficient was estimated at 4pp. Similarly, in Column 5, impacts on functioning toilets at RA1 are estimated at 5pp, while in Table 3.2 these were of 3pp. Interestingly, we find large significant impacts on ownership of functioning toilets even at RA2. As in our original specification, however, Column 6 shows that no significant impacts are observed at RA1 or RA2 on ownership of improved toilets.

What is driving these larger impacts seen in Table 4.4 compared with those from Table 3.2? As we can see from Figure 4.3(c), spillover households experience the same increase in toilet coverage as CLTS households. The inclusion of these in the control group reduces the programme effects because it increases average coverage rates in the control group, and this is assumed not to be driven by the programme. By removing spillover households from the control sample, and including these in the control (+ spillovers) group, we are gathering all the households we believe have been affected by the CLTS intervention in the same group, and comparing their outcomes with those of areas that did not receive CLTS or, we assume, its spillover effects. These pure control areas experience a much slower increase in toilet coverage, as seen in Figure 4.3. This increases our estimates of CLTS programme impacts, which compare treatment and control means. Our findings support the assumption that, by ignoring the possibility of programme spillovers, our estimates from Table 3.2 provide lower bound estimates of programme impacts. Nonetheless, these estimates must be interpreted with caution because, as already mentioned above, treatment and control samples in this case are markedly

different. The impacts estimated might not be caused by CLTS, but by our CLTS (+ spillovers) and pure controls samples actually having different toilet coverage trends, independent of the treatment.

Table 4.4. CLTS impact: CLTS (+ spillovers) versus pure controls.

Dependent var.: Condition of toilet	Constr./finished		Functioning		Improved	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated (γ)	0.05* (0.03)		0.05** (0.02)		0.04* (0.02)	
Treated $\times$ RA1 (γ_1)		0.07** (0.03)		0.05* (0.03)		0.03 (0.02)
Treated $\times$ RA2 (γ_2)		0.02 (0.03)		0.06** (0.03)		0.04 (0.02)
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Settlement population	Yes	Yes	Yes	Yes	Yes	Yes
LGA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Survey round fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (BL)	0.36	0.36	0.35	0.35	0.30	0.30
F-test $\gamma_1 = \gamma_2$ (p -value)		0.11		0.08		0.16
Number of settlements	104	104	104	104	104	104
Number of households	4,508	4,508	4,508	4,508	4,508	4,508
Number of observations	9,016	9,016	9,016	9,016	9,016	9,016

Note: Controls include age, age squared, gender, employment status and education attainment of household head, as well as a dummy variable indicating farming as the household's main economic activity. All controls are measured at baseline. Errors in parentheses are clustered at the unit of randomisation (TU). Stars indicate statistical significance: * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$.

In order to provide a robustness check to our estimation above, Table 4.5 presents the results of our second approach, in which we only compare settlements with either 100% CLTS households or 100% controls. As seen above, these two samples are, on average, identical observationally, except for some minor imbalances that can be controlled for. There is less of a concern about these two samples (pure CLTS and pure controls) having different sanitation trends, given that the balance is present also at the village and settlement levels. The improvement in balance across these two groups comes at a cost: because we have dropped all observations from settlements with intermediate shares of CLTS, we are left with a sample that is approximately 35% the size of our initial sample. This reduces our capacity to detect effects by significantly reducing the statistical power of our regression analysis. Therefore, it is not surprising that, as seen in Table 4.5, no statistically significant estimates are found. Nonetheless, point estimates for CLTS impacts on all three outcomes are in line with what we have found in our original approach from Table 3.2, and somewhat smaller than those presented in Table 4.4. This reinforces the caution with which estimates from Table 4.4 should be interpreted: while spillovers might be present and might be driving our impact estimates down, a comparison between treatment and control areas that should have experienced no spillovers does not suggest much larger impacts.

Two important conclusions can be taken from the results discussed in this section. First, toilet coverage trends over our study period suggest that spillovers at the community level are likely to exist and should be taken into consideration when estimating the

impacts of the CLTS intervention. Indeed, when we account for these, we see that CLTS impacts on ownership of functioning toilets increase, for our pooled estimates, for example, from 3pp to 5pp.¹² Secondly, these spillovers seem to magnify the impact of the intervention. In the first specification used in this section, the treatment group included both CLTS households and spillover households, while in the second specification, only pure CLTS households were included. If we assume that the CLTS has a higher impact on households who attend the triggering than on households who only hear about it, we should expect the estimated impacts from the first specification to be lower than those from the second. However, comparing the results from Column 4 in Tables 4.4 and 4.5, we see the opposite case: restricting our sample to settlements where no spillovers were possible leads us to lower, not higher, programme impacts. This could mean that spillovers act as a reinforcing, snowball effect, and could have important implications for future sanitation policy.

Table 4.5. CLTS impact: pure CLTS versus pure controls.

Dependent var.: Condition of toilet	Constr./finished		Functioning		Improved	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated (γ)	0.05*		0.05**		0.04*	
	(0.03)		(0.02)		(0.02)	
Treated $\times$ RA1 (γ_1)		0.07**		0.05*		0.03
		(0.03)		(0.03)		(0.02)
Treated $\times$ RA2 (γ_2)		0.02		0.06**		0.04
		(0.03)		(0.03)		(0.02)
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
LGA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Survey round fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (BL)	0.37	0.37	0.36	0.36	0.31	0.31
F-test $\gamma_1 = \gamma_2$ (p-value)		0.44		0.32		0.43
Number of settlements	82	82	82	82	82	82
Number of households	1,579	1,579	1,579	1,579	1,579	1,579
Number of observations	3,158	3,158	3,158	3,158	3,158	3,158

Note: Controls include age, age squared, gender, employment status and education attainment of household head, as well as a dummy variable indicating farming as the household's main economic activity. All controls are measured at baseline. Errors in parentheses are clustered at the unit of randomisation (TU). Stars indicate statistical significance: * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$.

In a future research paper, we will explore further the nature of these CLTS programme spillovers, and we will study their transmission channels to try to understand how the information flows from CLTS to spillover areas. This has yet to be analysed more extensively to provide more robust conclusions, and it is beyond the scope of this report.

¹² Coefficients for 'Treated $\times$ RA1' and 'Treated $\times$ RA2' for our original estimation were compared to those from specifications 1 and 2 that account for spillovers, and we cannot reject the hypothesis that the coefficients are, in fact, identical.

5. Conclusions and Further Research

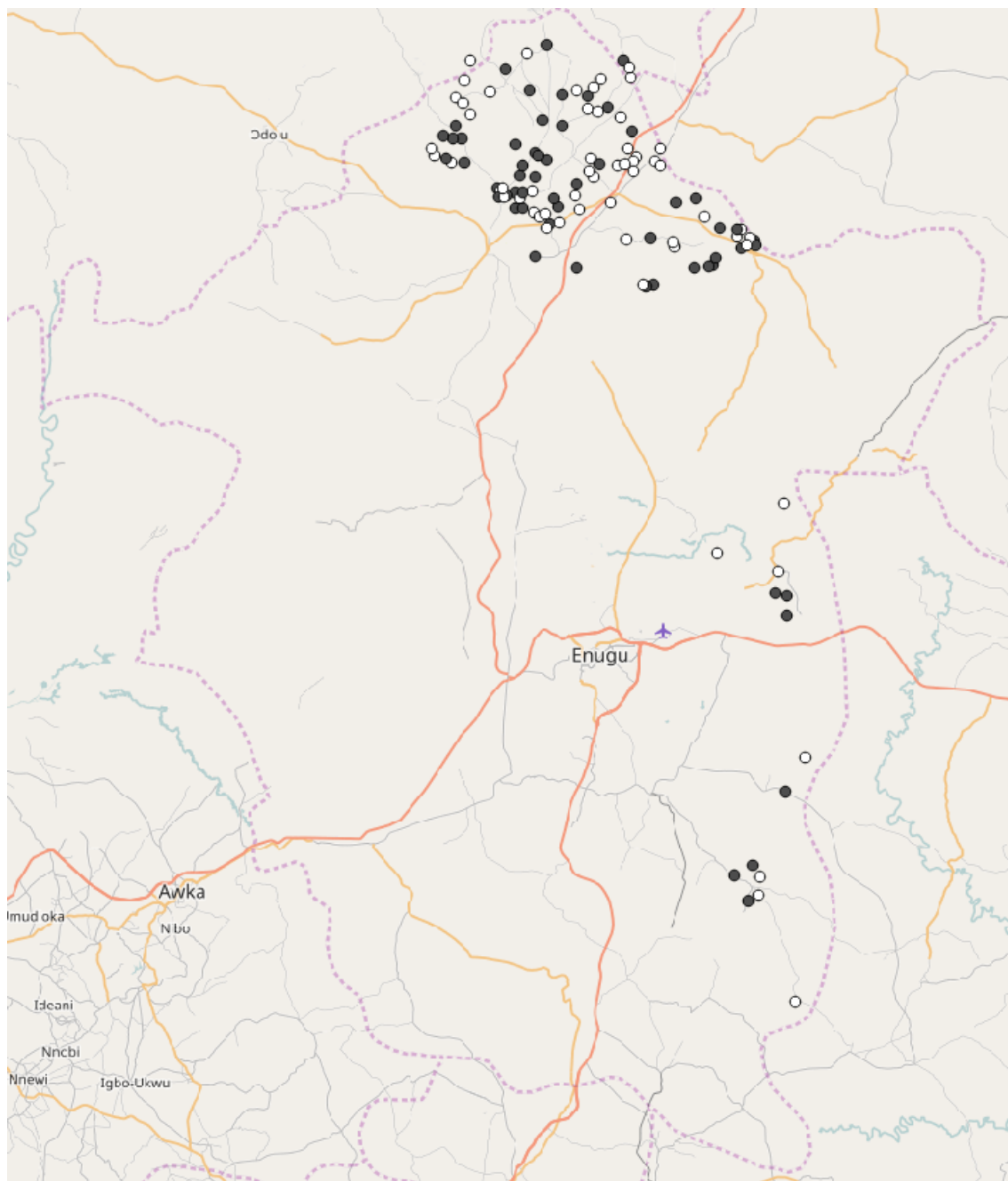
Using household data collected over two years following the implementation of the CLTS intervention, we show that the positive (negative) effect of CLTS on toilet ownership (OD) detected six to twelve months after the implementation are no longer statistically significant. This seems to be driven by households in control areas catching up with households in treatment areas between the two follow-up survey rounds.

These short-term impacts of CLTS are coming from a specific group of households that could be thought of as the most vulnerable: household heads have no education at all (primary school not completed), they have children and senior members (making them more exposed to disease) and they are relatively less wealthy than the rest of the households in the sample. Consequently, they are also less likely to have access to financial instruments (less likely to have savings and debt). The reason why CLTS affects this segment of the population in particular goes beyond the scope of this report, and this will be explored in greater detail in a separate research paper.

The fact that no impact is observed after two years of implementation, because of households in control areas catching up, led us to explore the possibility of spillovers from CLTS. CLTS has an important information component and there are reasons to think that information may have flowed from CLTS to control areas, even though the design of treatment units was intended to avoid this. Initial evidence is consistent with the existence of spillovers and it suggests that the estimated impacts may be a lower bound and that there could be some more longer-term impacts of CLTS that have so far been undetected. Further work is needed to define a more robust methodology to estimate CLTS impact effects in the presence of spillovers, given the evaluation design pursued in this study.

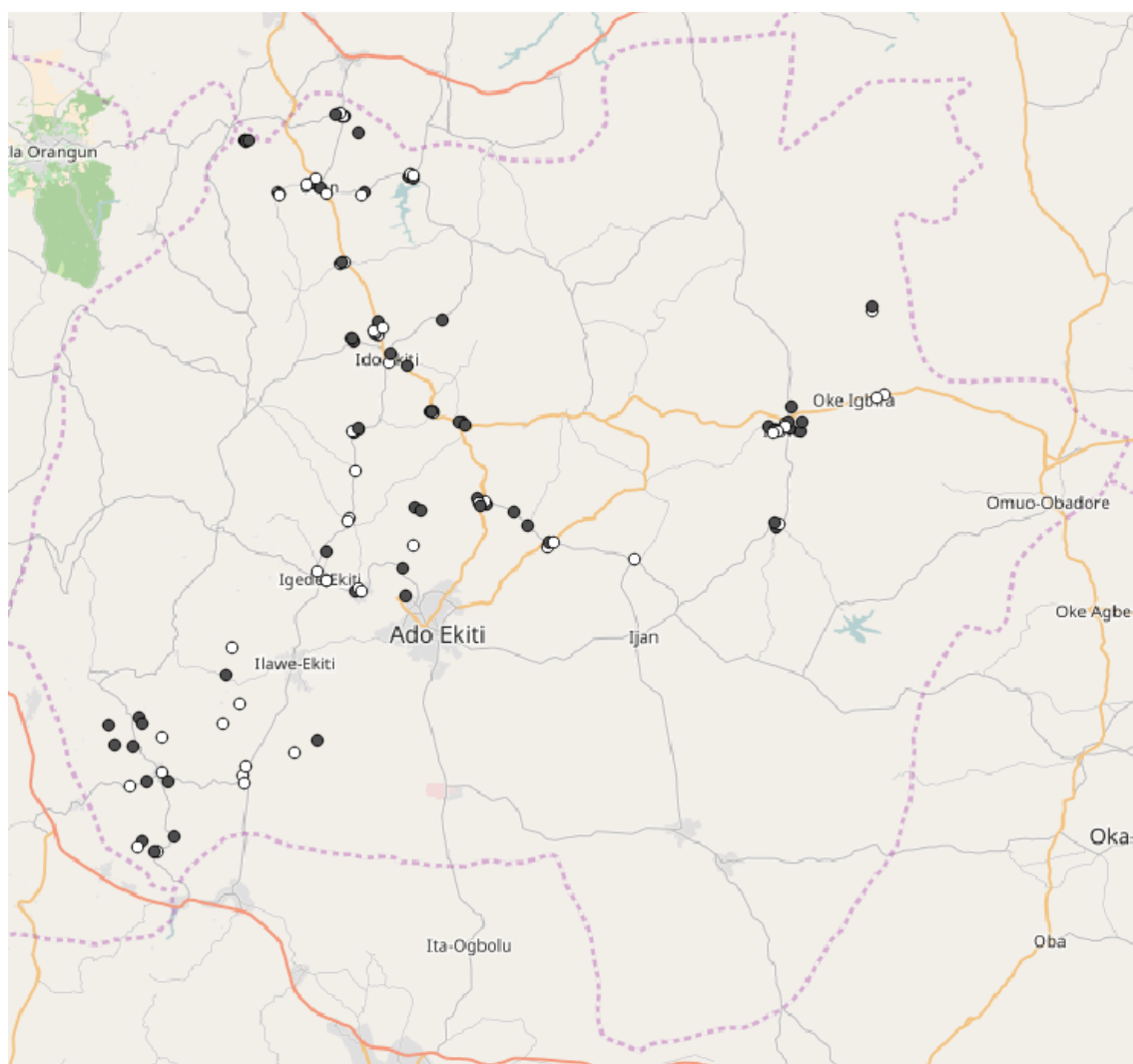
Appendix

Figure A.1. Location of CLTS and control TUs, in Enugu state.



Note: CLTS (black) and control (white) TUs in the state of Enugu. Locations indicate the centroid of a polygon formed by all of the households interviewed in each TU.

Figure A.2. Location of CLTS and control TUs, in Ekiti state.



Note: CLTS (black) and control (white) TUs in the state of Ekiti. Locations indicate the centroid of a polygon formed by all of the households interviewed in each TU.

References

- Abramovsky, L., Augsburg, B. and Oteiza, F. (2015), 'Sustainable Total Sanitation – Nigeria: Baseline Report', Institute for Fiscal Studies (IFS) mimeo (https://www.ifs.org.uk/uploads/publications/mimeos/Abramovsky_et_al_Baseline%20report%20Nigeria%20STS.pdf).
- Abramovsky, L., Augsburg, B. and Oteiza, F. (2016), 'Sustainable Total Sanitation – Nigeria: Rapid Assessment Report', Institute for Fiscal Studies (IFS).
- Baird, S., Bohren, J. Aislinn, McIntosh, C. and Özler, B. (2016), 'Optimal Design of Experiments in the Presence of Interference', University of Pennsylvania Working Paper no. 16-025.
- Cameron, L. A, Olivia, S. and Shah, M. (2015), 'Initial Conditions Matter: Social Capital and Participatory Development', IZA Discussion Paper no. 9563, available at: <http://ftp.iza.org/dp9563.pdf>.
- Duflo, E., Glennerster, R. and Kremer, M. (2008), 'Using randomization in development economics research: a toolkit', in T. P. Schultz and J. A. Strauss (eds), *Handbook of Development Economics*, Volume 4, Chapter 61, Amsterdam: North-Holland. (We refer to the version posted in 2007 by the Centre for Economic Policy Research, CEPR Discussion Paper no. 6059, http://cepr.org/active/publications/discussion_papers/dp.php?dpno=6059.)
- Hudgens, M. G. and Halloran, M. E. (2008), 'Toward causal inference with interference', *Journal of the American Statistical Association*, vol. 103, pp. 832–842.
- McKenzie, D. (2012), 'Beyond baseline and follow-up: the case for more T in experiments', *Journal of Development Economics*, vol. 99, pp. 210–221 (<https://doi.org/10.1016/j.jdeveco.2012.01.002>).
- Özler, B. (2016), 'Definitions in RCTs with interference', Technical report, *Development Impact*, New York: World Bank.
- UNICEF and WHO (2015), 'Progress on sanitation and drinking water – 2015 update and MDG assessment'.