



USAID
FROM THE AMERICAN PEOPLE



Baseline Brief for an Impact Evaluation of Governance, Accountability, Participation and Performance (GAPP): SMS for Better Service Provision in Uganda

March 2015

This publication was produced at the request of the United States Agency for International Development. It was prepared independently by Jon Helfers (Social Impact), Mitchell Goist (Pennsylvania State University and Social Impact), Guy Grossman (University of Pennsylvania and Social Impact), Melina Platas Izama (Stanford University and Social Impact), and Jonathan Rodden (Stanford University and Social Impact).

Baseline Brief for an Impact Evaluation of Governance, Accountability, Participation and Performance (GAPP): SMS for Better Service Provision in Uganda

March 23, 2015

AID-OAA-M-13-00011

DISCLAIMER

The author's views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

CONTENTS

Acronyms	2
Executive summary.....	3
Introduction to the intervention and evaluation design	5
School audit	6
Health center audit.....	12
Balance between treatment and control groups	16
Annex 1: Balance tests.....	20
Annex 2: Number of missing variables	22

ACRONYMS

DEO	District Education Officer
DHO	District Health Officer
GAPP	Governance, Accountability, Participation and Performance
HC	Health Center
RCT	Randomized Control Trial
RTI	RTI International
USAID	U.S. Agency for International Development

EXECUTIVE SUMMARY

This report summarizes the findings of baseline data collection measuring the quality of government service provision in Arua District. This randomized control trial (RCT) was designed to evaluate the impact of the U-Bridge project, a SMS-based system designed to foster accountability by linking citizens, service providers and government officials.

The U-Bridge initiative, developed in partnership with UNICEF Uganda, allows citizens to send SMS messages at no cost to district officials, detailing their location and issue with service provision. This has the potential to improve government service provision by providing common information between citizens and government officials. Unlike past informational interventions, the U-Bridge system allows for two-way communication, so district officials and the research team can also contact citizens in order to ask questions or provide information about particular aspects of government services. In order to measure the impact this has on local public goods provision, this baseline survey measures the quality of existing services in the health and education sectors.

In order to evaluate the program's effectiveness, the research team began by specifying clusters of Local Councils, the finest grain administrative unit in Uganda, around different health centers (HCs). The clusters were then divided into two groups (treatment and control) within Arua, and one quasi-control group from areas in districts bordering Arua. Recognizing variation in the size and capacity of the HCs, randomization was stratified between two types of HCs: the smaller Level II HCs and the larger Level III HCs. Both the treatment and control groups consists of 11 Level II HCs and 13 Level III HCs each, for a total of 23 treatment and 24 control clusters. In addition, there are 26 quasi-control HCs selected from the areas bordering Arua. The quasi-control clusters will allow the research team to estimate spillover effects from the treatment areas.

In each of the treatment and control clusters, the evaluation team oversaw a series of HC and primary school audits to measure the quality of health and education service provision. The goal of the baseline survey was to establish the control group as a credible counterfactual for the SMS intervention treatment, and construct a pre-test measure to estimate the effect of the treatment.

For schools, the primary outcome indicators are: teacher absenteeism; pupil absenteeism; evidence of teaching; student performance; necessary school materials; and management of the school. For HCs, the primary outcome indicators are: worker absenteeism; drug stockouts; waiting time; necessary equipment; and management of the clinic. Collection of the baseline survey data began on 26 June 2014 and was completed on 8 November 2014.

In the majority of schools, the teacher was present in both observed classrooms, but there were numerous instances of teacher absence. Most teachers in the sample were inexperienced, with all but a few having less than five years of teaching experience.

There are also considerable deficiencies in the materials available to students. A vast majority of students did not have uniforms, half of students had books, and slightly less than half had writing utensils. The average attendance across classrooms was around 50%. Most of the classrooms are covered and have blackboards—although significantly more have blackboards. Most of the schools surveyed do not have a functioning toilet or permanent houses for teachers. Most schools have a total of seven streams (or grade levels), with a minimum of 3 and a maximum of 14. There is considerable variation in the total number of pupils at each school, with the mean slightly over 800, but ranging from 300 to almost 2000.

One of the key aspects of this intervention that distinguishes it from previous SMS-based treatments is

that it allows two-way communication between district officials and citizens. Therefore, it is critical to know the extent to which schools had access to officials and other top-down monitoring mechanisms before the intervention. Most schools have limited interactions with the District Education Officer (DEO)—with the modal category being one phone call and zero visits per month. If the intervention is successful, we would expect to see an increase in visits and phone calls from the DEO.

In addition to the school audit, a health center audit was administered in 74 total health centers: 47 from Arua (treatment and control clusters), and 26 from quasi-control clusters; specifically, 9 from Maracha; 7 from Nebbi; 3 from Zombo; and 7 from Yumbe. As mentioned above, treatment and control groups were balanced between the two levels of health centers included in the study, HC Level II and Level III. Outcomes for HCs can be broken down into the following five groups: worker absenteeism; drug stockouts; patient waiting time; equipment available; and management of the clinic.

In a majority of HCs, nearly 75% of workers were absent without an excuse. Approximately 12% of all HCs in the sample were not even open when the audit took place during operating hours. Further, only one quarter of surveyed HCs had the required medical units and nearly one quarter of facilities did not have electricity. On the positive side, almost all HCs had access to sterilized equipment and clean water, and there were functioning toilets in all of the surveyed HCs.

Balance checks were performed to test the comparability of the treatment and control groups. Previous research has suggested that any differences greater than one-fourth of a standard deviation should be concerning. For the HC dataset, it seems that this criteria is met for absenteeism without an excuse, health inspector visits, sterilization, and DHO visits. For the school audit, only student absenteeism crosses the one-quarter standard deviation threshold. These variables will be controlled for in future analyses.

INTRODUCTION TO THE INTERVENTION AND EVALUATION DESIGN

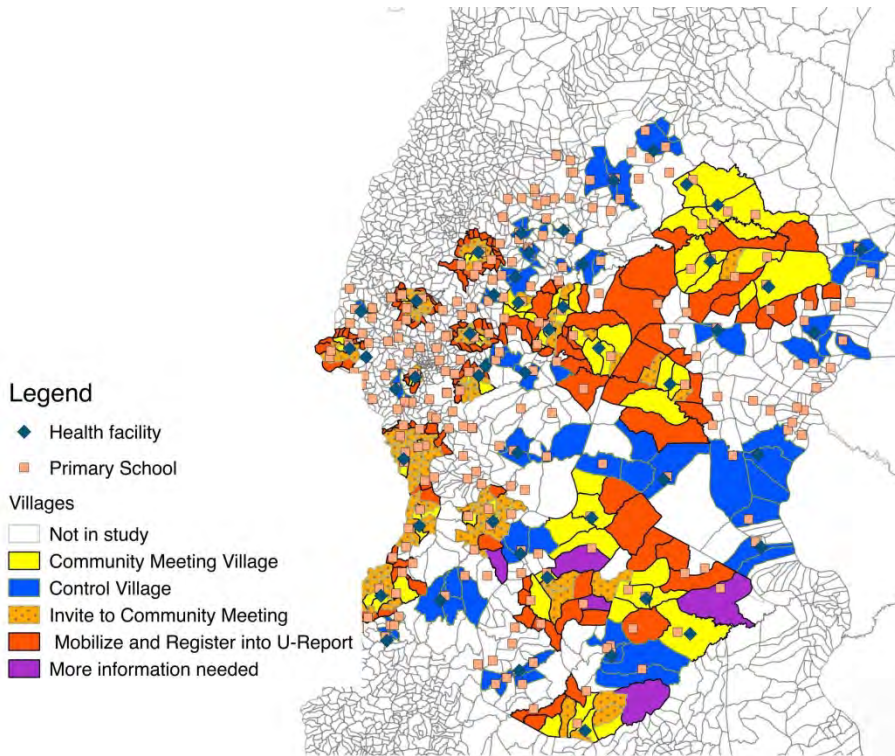
This report summarizes the findings of baseline data collection measuring the quality of government service provision in Arua District. This randomized control trial (RCT) was designed to evaluate the impact of the U-Bridge project, a SMS-based system designed to foster accountability by linking citizens, service providers and government officials. The U-Bridge project is funded USAID, and implemented by RTI International's Governance, Accountability, Participation, and Performance (GAPP). The evaluation is administered through Social Impact. Hatchile Consulting, Ltd. undertook the initial data collection presented here. The principal investigators on the project are Guy Grossman (University of Pennsylvania) and Jonathan Rodden (Stanford University), who are being assisted by Melina Platas Izama, research assistant, and Jon Helfers, in-country research coordinator.

The U-Bridge initiative, developed in partnership with UNICEF Uganda, allows citizens to send SMS messages, at no cost, to district officials, detailing their location and issue with service provision. This has the potential to improve government service provision by providing common information between citizens and government officials. Unlike past informational interventions, the U-Bridge system allows for two-way communication, so district officials and the research team can also contact citizens in order to ask questions or provide information about particular aspects of government services. In order to measure the impact this has on local public goods provision, this baseline survey measures the quality of existing services in the health and education sectors.

In order to evaluate the program's effectiveness, the research team began by specifying clusters of Local Councils, the finest grain administrative unit in Uganda, around different health centers (HCs). The clusters were then divided into two groups (treatment and control) within Arua, and one quasi-control group from areas in districts bordering Arua. Recognizing variation in the size and capacity of the HCs, randomization was stratified between two types of HCs: the smaller Level II HCs and the larger Level III HCs. Both the treatment and control groups consists of 11 Level II HCs and 13 Level III HCs each, for a total of 23 treatment and 24 control clusters. In addition, there are 25 quasi-control HCs selected from the areas bordering Arua. The quasi-control clusters will allow the research team to estimate spillover effects from the treatment areas.

Figure I shows the treatment and control clusters in Arua district. Control communities serviced by HCs are presented in blue and treatment communities in yellow, orange and red. Residents in treatment communities were invited to participate in the U-Bridge initiative through either community meetings or a door-to-door registration efforts. The figure includes 47 HCs and 84 primary schools in both treatment and control areas. Quasi-control areas are not included in the figure below.

Figure 1: Treatment and Control Clusters in Arua



In each of the treatment and control cluster, the evaluation team oversaw a series of HC and primary school audits to measure the quality of health and education service provision. The goal of the baseline survey was to establish the control group as a credible counterfactual for the SMS intervention treatment, and construct a pre-test measure to estimate the effect of the treatment. These audits represent the first of the unannounced audits. According to the original design, additional audits were to occur after 6, 9, and 12 months of the start date of the program. Due to some initial challenges in program start-up, this timeline for future data collection is under review. For schools, the primary outcome indicators are: teacher absenteeism; pupil absenteeism; evidence of teaching; student performance; necessary school materials; and management of the school. For HCs, the primary outcome indicators are: worker absenteeism; drug stockouts; waiting time; necessary equipment; and management of the clinic. Collection of the baseline survey data began on 26 June 2014 and was completed on 8 November 2014. Data cleaning and management continued through January 2015. Findings from these school audits and health center audits are presented in the sections below.

SCHOOL AUDIT

As stated above, the school audit consisted of unannounced spot checks by enumerators. These covered a total of 134 primary schools, 50 of which are in the quasi-control group, and the remaining 84 divided evenly between treatment and control. The audit survey is divided into three sections. The first section includes data on conditions of the school overall. These include the number of functioning toilets, number of classrooms, temporary and permanent housing for instructors, and other variables that reflect the overall management and infrastructure available in the school. The second section includes variables on specific classrooms. Enumerators were instructed to visit two primary level three

(P3) classrooms. If there was only one P3 class at the particular school, enumerators visited a P3 and P4 classroom. This second section includes information on the resources available to students (such as adequate uniforms, books, and writing utensils) and evidence of teaching in that particular classroom. In addition, enumerators interviewed the teachers of these classes, measuring their experience (amount of time, in days, as a teacher) and education (measured in years of schooling). Finally, in the third section, enumerators conducted interviews with the head teacher of each school, collecting information on number of planning meetings, staff absenteeism, and other school-wide variables.

Unfortunately, during the course of the audit, data on the second classroom visited (whether P3 or P4) was lost when transporting the paper surveys from Arua to Kampala. The amount of missing information varies by question; Figure 2 displays the number of missing observations as an annotation in the upper left-hand corner. This particularly affected the teacher interview variables from the second classroom visited. Detailed descriptions, and a breakdown of extent of missingness by school, is illustrated in Annex 2. Reassuringly, the missing schools do not seem to vary systematically, either in their distribution between treatment and control groups, or in the distributions of the variables themselves.

Figure 2 illustrates the following classroom level outcomes for all 84 schools in both the control and treatment groups: whether a teacher is present; proportion of students with uniforms; proportion of students with books; proportion of students with writing utensils; ratio between number of students present and the total students in the class; the tenure of the teacher instructing the class (measured in days); and that teacher's education (measured in years). Figure 2 also presents the number of classrooms missing. The light red shade indicates full information (two classrooms, either two P3 or one P3 and one P4), the light blue shade indicates information calculated from one classroom (almost always the first P3 class), and finally the number of complete missing (no classrooms surveyed) is annotated in red in the upper-left corner of the graph. The information presented is aggregated across treatment and control groups. Differences between these groups are discussed later in the brief.

Figure 2: Classroom Averages for School Outcomes



As shown in the upper, left-hand figure (q31), in a plurality of schools visited, the teacher was present in both classrooms. However, while this is the modal outcome, there were numerous instances of teacher absences. In 34 schools a teacher was present in one classroom but not the other (indicated by “split”),

and in 33 schools no teacher was present in surveyed classrooms. Data on teacher absences was not collected in eight of the 134 schools.

Of those classrooms in which teachers were present, the education (q48) and teacher tenure (q47) varies significantly. Most teachers in the sample were inexperienced, with all but a few having less than five years of teaching experience. Also, almost all of the teachers had less than five years of education (q48) and the median education was two years. Missing values are higher for teacher education and tenure because of teacher absenteeism in the audited schools.

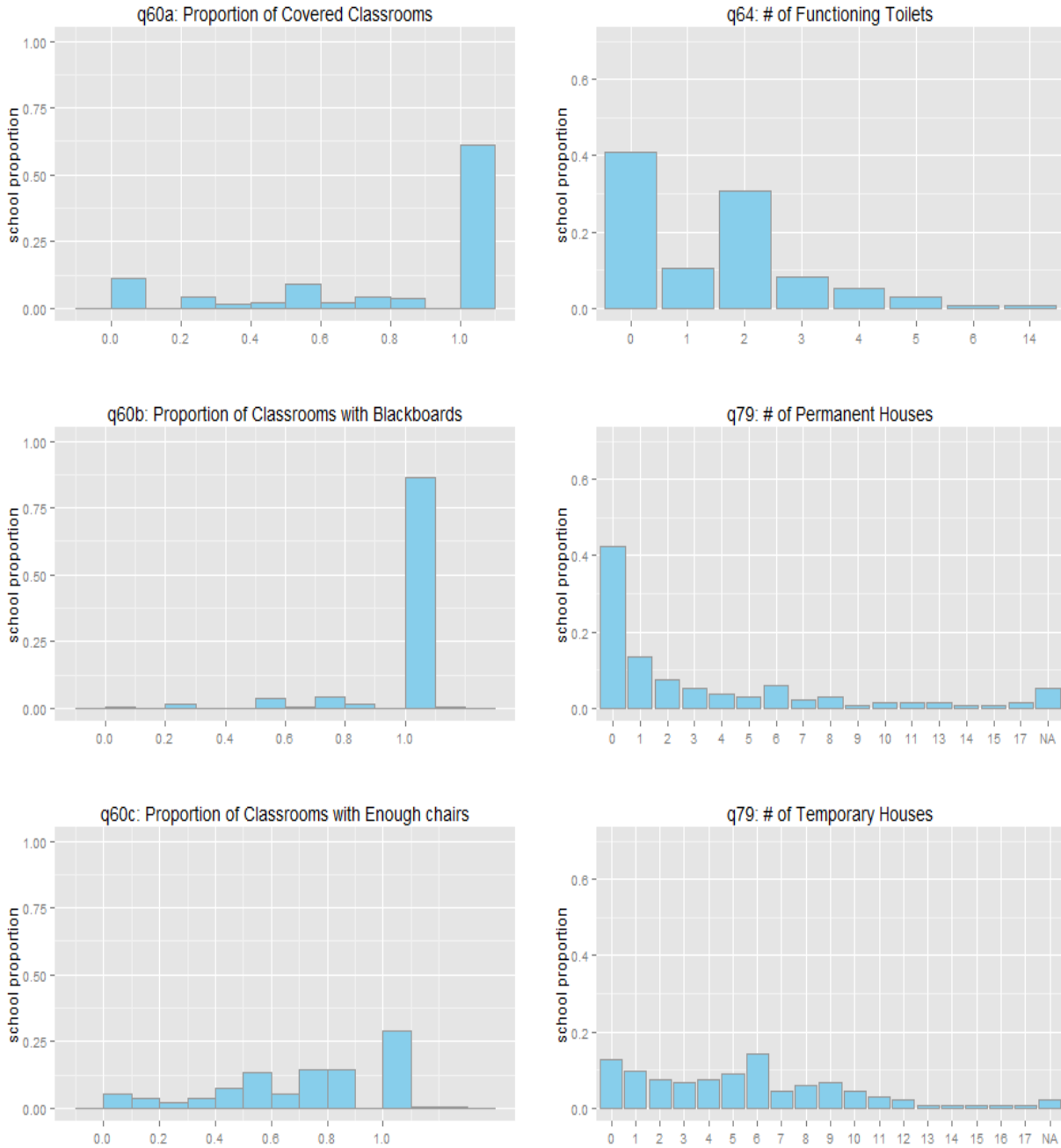
There are also considerable deficiencies in the materials available to students. A vast majority of students did not have uniforms (q42), with uniforms defined as the required tops and bottoms. In the median school, only 28% of students had uniforms. Remaining with Figure 2, the situation is slightly improved when looking at the proportion of students with books (q43) or writing utensils (q44). On average, half of students in a given Arua school have books, while this figure is slightly less for writing utensils. There was considerable variation in the availability of materials across classes. There were schools where no student had a book or writing utensil, and schools where every student had a book and writing implement.

Student attendance exhibited similar variation. The average attendance across classrooms was around 50%. If the intervention is effective and citizen complaints through the SMS system create pressure on public officials and increase officials' awareness of problems at the schools, then we would expect to see improvement across these indicators.

SCHOOL-LEVEL OUTCOMES

Figure 3 illustrates outcomes at the school level. Most of the classrooms are covered and have blackboards—although significantly more have blackboards. Most of the schools surveyed do not have a functioning toilet or permanent houses for teachers. The number of temporary houses for teachers varies considerably across schools.

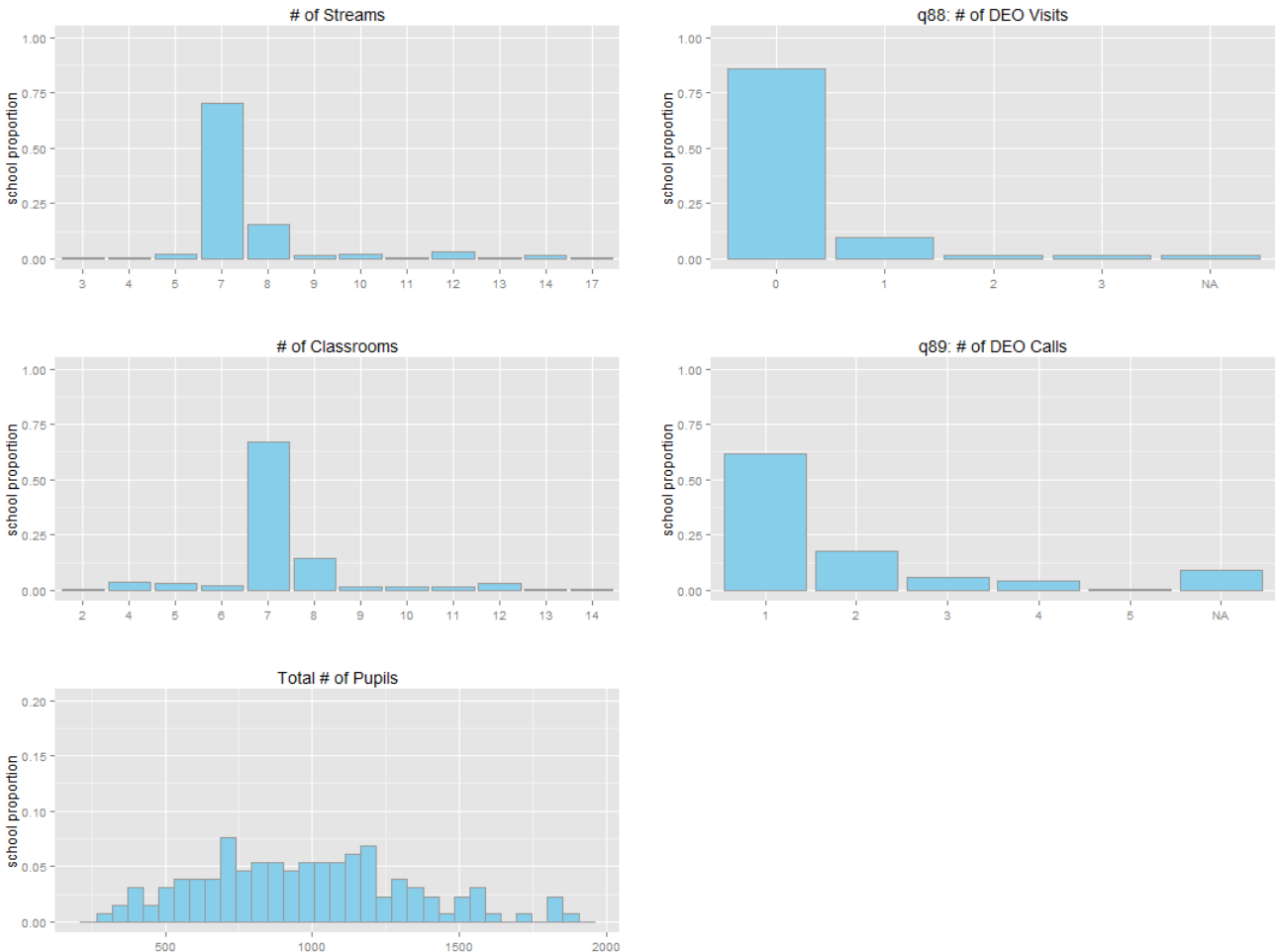
Figure 3: School Outcomes



Potential covariates, or other variables that could impact the extent of service provision, are illustrated in Figure 4. For example, District officials might be more responsive to the needs of larger schools. Most schools have a total of seven streams (or grade levels), with a minimum of 3 and a maximum of 14. Not surprisingly, the number of classrooms closely follows the distribution of the number of streams. There is considerable variation in the total number of pupils at each school, with the mean slightly over 800, but ranging from 300 to almost 2000.

One of the key aspects of this intervention that distinguishes it from previous SMS-based treatments is that it allows two-way communication between district officials and citizens. Therefore, it is critical to know the extent to which schools had access to officials and other top-down monitoring mechanisms before the intervention. As such, Figure 5 also presents information on the number of interactions with the District Education Officer (DEO). Most schools have limited interactions with the DEO—with the modal category being one phone call and zero visits per month. If the intervention is successful, we would expect to see an increase in visits and phone calls from the DEO.

Figure 4: School Covariates



HEALTH CENTER AUDIT

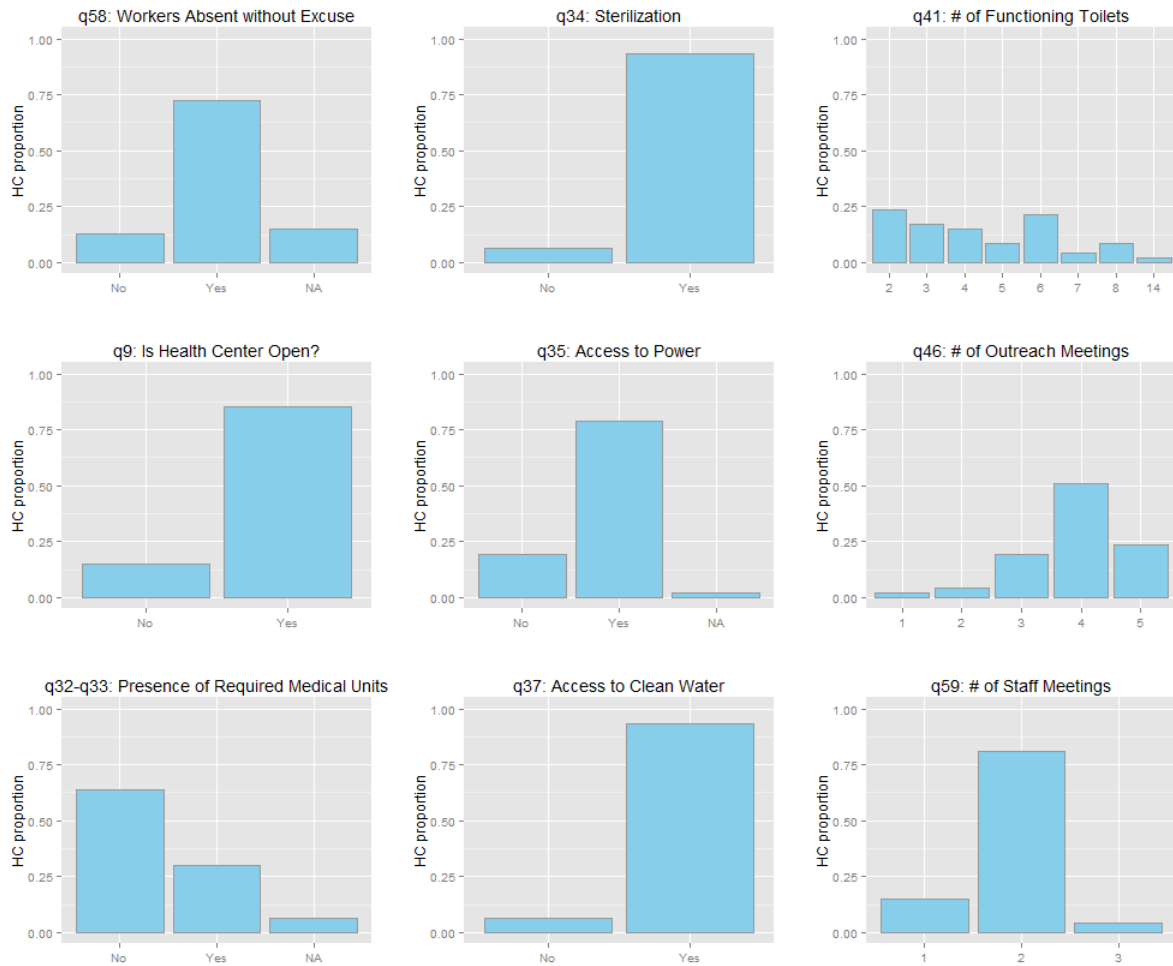
In addition to the school audit, a health center audit was administered in 74 total health centers: 47 from Arua (treatment and control clusters), and 26 from quasi-control clusters; specifically, 9 from Maracha; 7 from Nebbi; 3 from Zombo; and 7 from Yumbe. As mentioned above, treatment and control groups were balanced between the two levels of health centers included in the study, HC Level II and Level III. Outcomes for HCs can be broken down into the following five groups: worker absenteeism; drug stockouts; patient waiting time; equipment available; and management of the clinic.

Figure 5 lists various outcomes for all HCs. In a majority of HCs, nearly 75% of workers were absent without an excuse. Approximately 12% of all HCs in the sample were not even open when the audit took place during operating hours. Further, only one quarter of surveyed HCs had the required medical units¹. Nearly one quarter of facilities did not have electricity.² On the positive side, almost all HCs had access to sterilized equipment and clean water. There were functioning toilets in all of the surveyed HCs, and the number of toilets ranged from 2 to 14. With respect to the management of clinics, most HCs reported holding 4 community outreach meetings and 2 staff meetings every month.

¹ For a Level II Health Center, the required medical unit is an outpatient facility. Level III Health Centers are required to have both an outpatient facility and a maternity ward.

² Specifically, these were calculated from the question “What is the main source of water (or power) for the facility? Respondents who had selected “No Power Supply”, were recorded as zero values.

Figure 5: Health Center Outcomes



The question of drug stockouts is investigated in more detail in Figure 6, which shows the number of days the clinic had been without each drug included in the list of essential drugs (antimalarials, deproprovera, suphadoxine, measles vaccine, oral rehydration salts (ORS), and cotrimoxazole) at the time of the audit. A value of zero indicates that the drug was currently in stock. At the time of the audit, 22% of HCs were out of antimalarials, 24% were out of deproprovera, 29% were out of suphadoxine, 31% were out of the measles vaccine, 29% were out of ORS, and 27% were out of Cotrimoxazole. Shortages were most acute for ORS, with stockouts ranging up to 80 days, with many HCs experiencing shortages greater than 25 days. Antimalarials faced a similar, although less severe problem, with many clinics experiencing shortages over 25 days. Supplies of cotrimoxazole and deproprovera were the least likely to be out of stock across all clinics.

Figure 6: Health Center Drug Stockouts

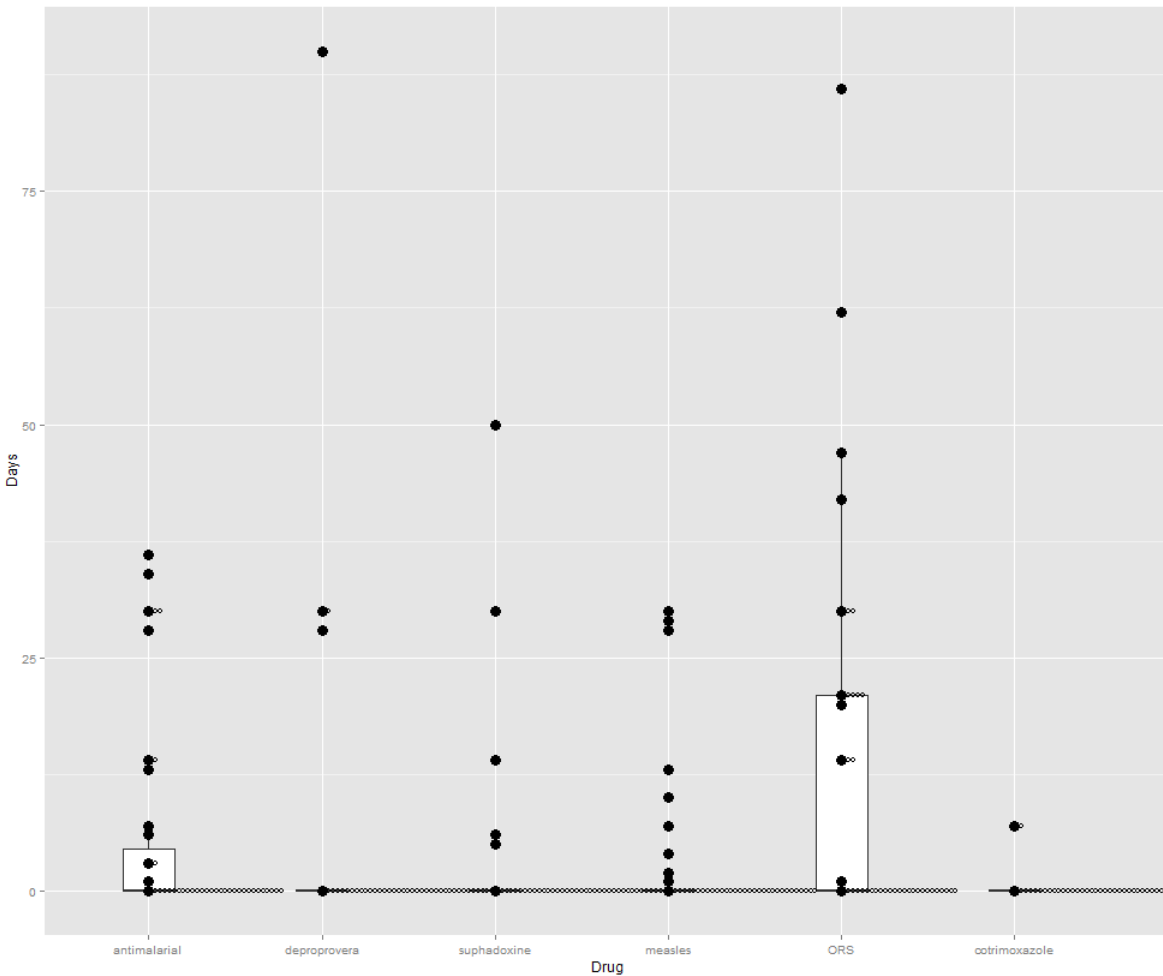
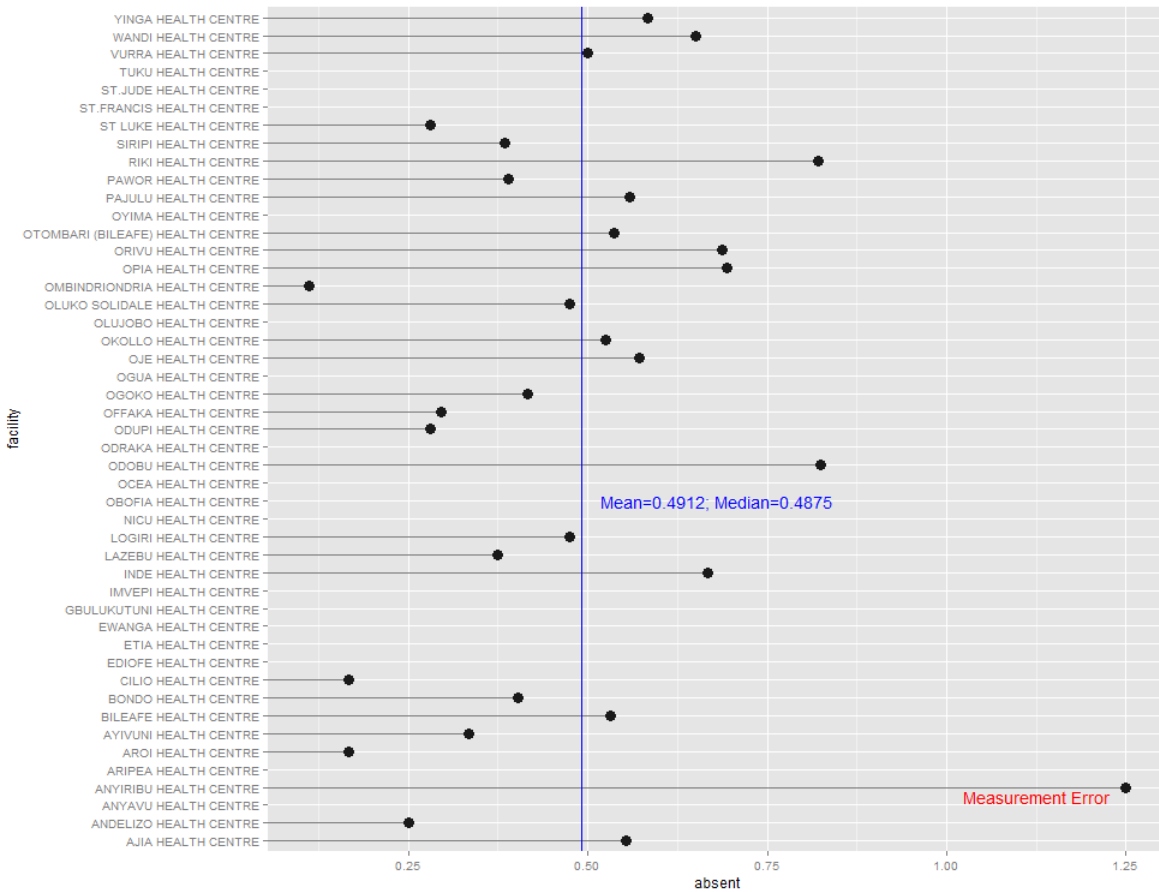


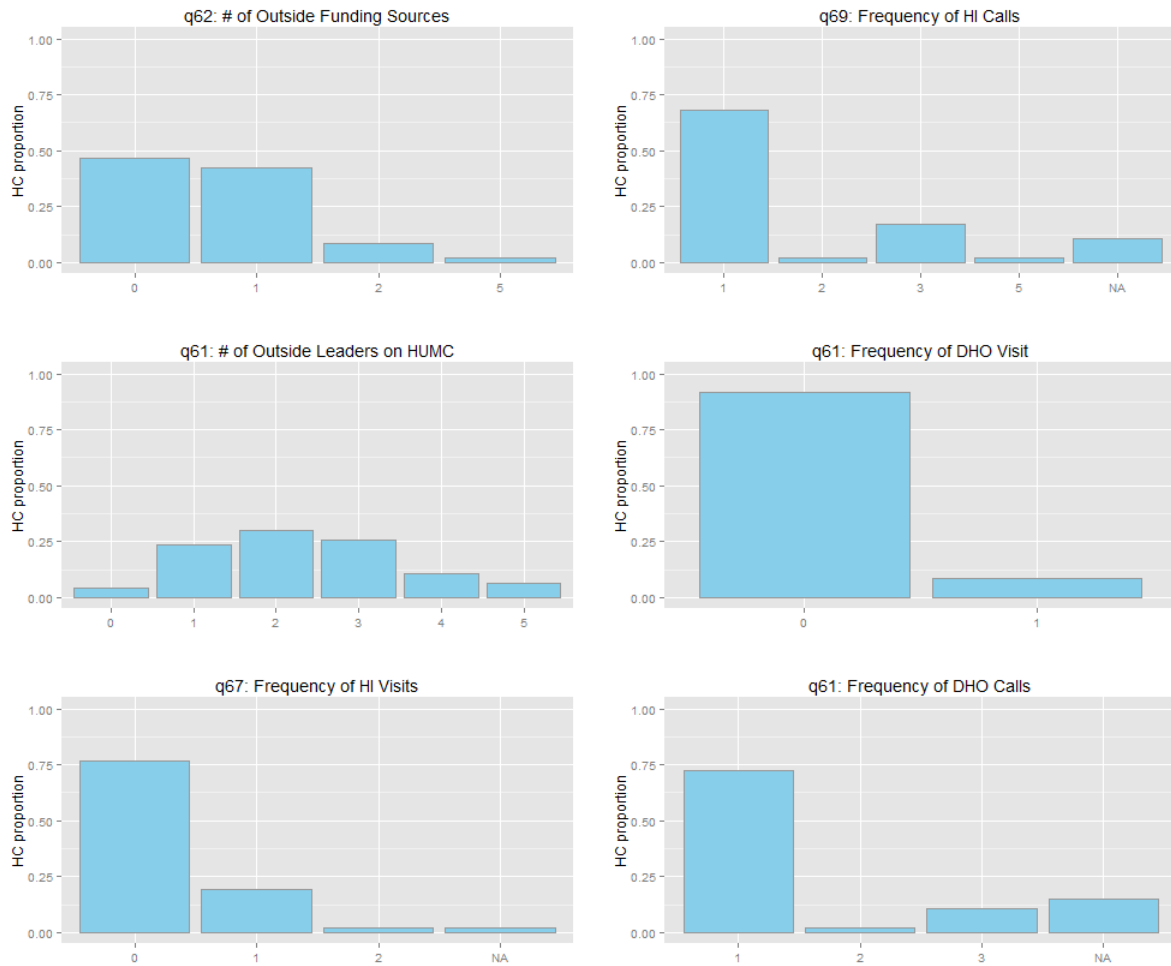
Figure 7 explores the question of health worker absenteeism in greater detail. This figure illustrates the percentage of workers who had not signed the attendance log over an average of four days, broken down by each HC in the treatment and control groups. It is important to note that many of these HCs have missing values, indicating either enumerator error or, more likely, that logs were not adequately maintained to gauge absenteeism. Of those with logs, 14 had over 50% of registered workers absent. However, the Anyiribu Health Center had more workers sign the attendance log than were listed as recognized staff. This is either due to measurement error, poor bookkeeping, or some attempt to mislead investigators.

Figure 7: Health Center Absenteeism



Finally, Figure 8 illustrates potential covariates, or variables that could mediate the impact of service provision, for the HCs surveyed. Most of the clinics surveyed did not have other outside funding sources. Most, however, had local leaders such as LC chairpersons or members of parliament on the Health Unit Management Council. All surveyed clinics had at least one call from a health inspector (HI) per month, although most did not receive any visits by the HI. Similarly, every clinic had at least one call from the District Health Officer (DHO), but almost none received a personal visit.

Figure 8: Health Center Covariates



BALANCE BETWEEN TREATMENT AND CONTROL GROUPS

Randomizing which health center clusters receive the SMS intervention allows the researchers to ensure that assignment to this treatment is not correlated with other unobserved variables that might also affect the level of service provision. However, in rare cases, such randomization does not always balance these potential confounders, and, by random chance, confounders are more concentrated in the treatment or control condition. While this should not be a concern given the randomization and power calculations done prior to the analysis, this section briefly examines whether treatment and control groups are balanced on relevant covariates.

Figure 9 shows density plots for the treatment (light red) and control (light blue) conditions across a variety of potential confounding variables. As can be seen from the figure, the distributions are relatively even—no single variable seems highly concentrated in one of the treatment conditions. Figure 10 supplies a similar picture for the health audit. Once again, according to these observable variables, there do not seem to be significant imbalances across either the treatment or control groups.

Figure 9: School Covariate Balance

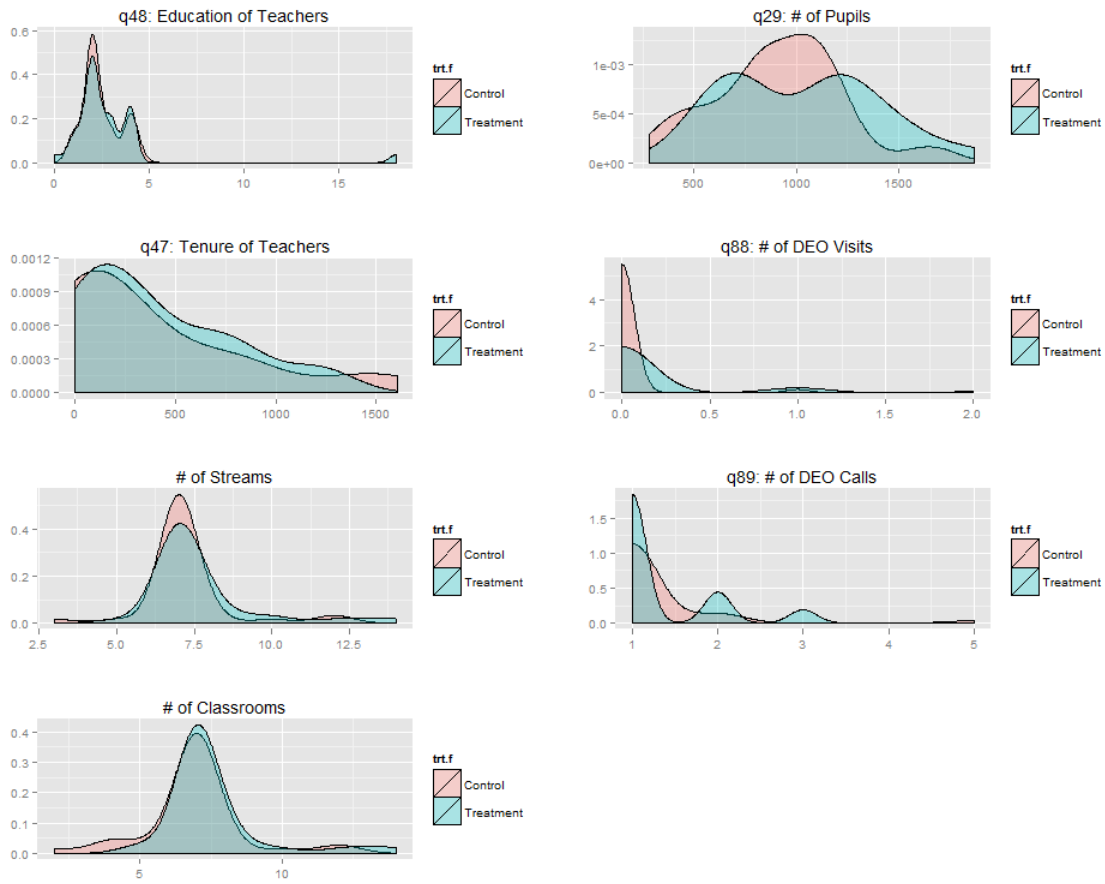
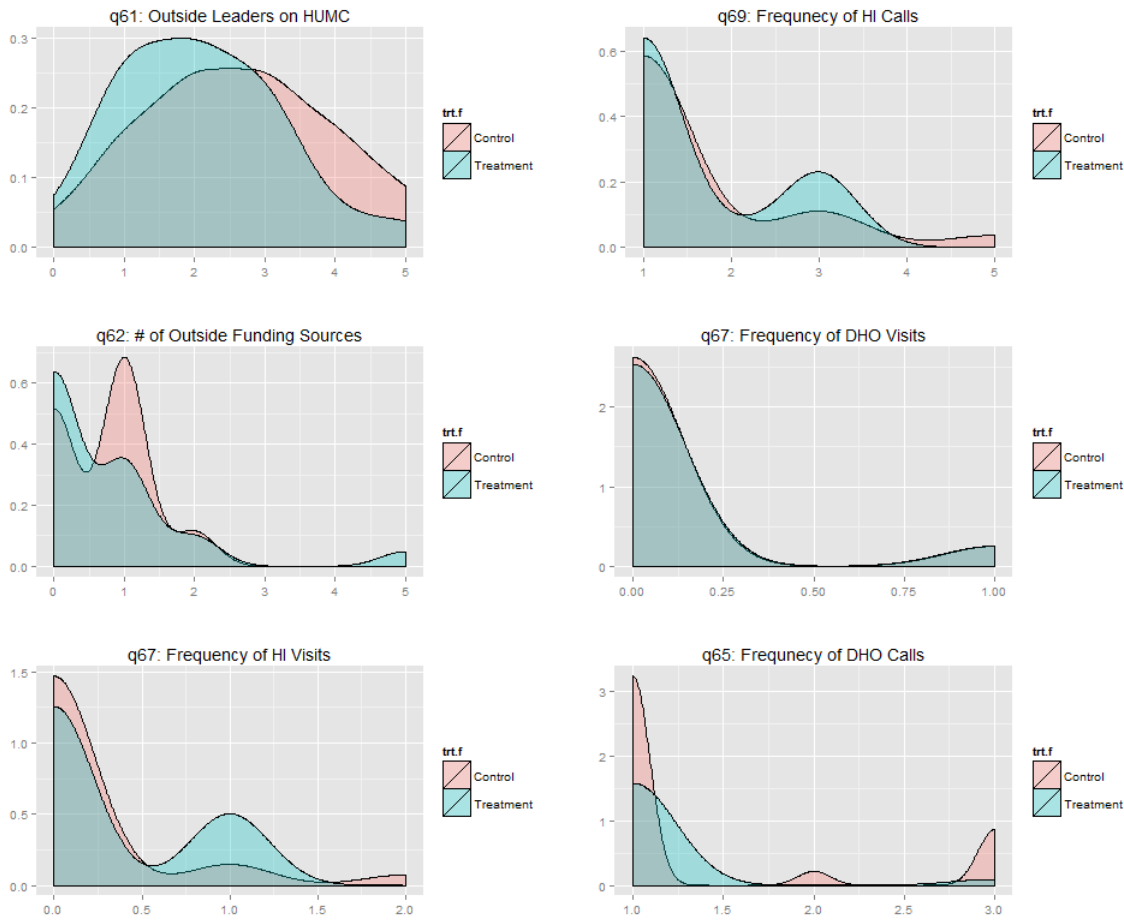


Figure 10: Health Center Covariate Balance



While the Figures 9 and 10 illustrate a visual confirmation of balance across treatment and control groups, balance is further investigated in Tables 1 and 2. Table 1 compares the treatment and control means, and the distance these means differ by with respect to standard deviation. Previous research has suggested that any differences greater than one-fourth of a standard deviation should be concerning. For the HC dataset, it seems that this criteria is met for absenteeism without an excuse, health inspector visits, sterilization, and DHO visits. Table 2 enumerates the same variables for the school audit. In this case, only student absenteeism crosses the one-quarter standard deviation threshold. These variables can be controlled for in the analysis. Appendix 2 more rigorously addresses questions of balance between treatment and control groups.

Table 1: Health Center Covariate Balance

	Treatment Mean	Control Mean	Difference	Std. Bias
q58: absent without excuse	0.92	0.81	0.11	0.40
q9: is health center open	0.85	0.94	-0.09	-0.24
q32: necessary facilities	0.23	0.31	-0.08	-0.19
q43: sterilization	0.92	1.00	-0.08	-0.28
q35: access to power	0.85	0.81	0.03	0.09
q37: clean water	1.00	0.94	0.06	-
q41: number of toilets	4.23	4.44	-0.21	-0.09
q46: outreach	3.69	4.06	-0.37	-0.49
q59: staff meetings	2.00	1.88	0.12	0.31
q62: funding	0.92	0.75	0.17	0.13
q61: HUMCS	2.46	2.44	0.02	0.02
q67: health inspector visit	0.31	0.12	0.18	0.38
q69: health inspector call	1.69	1.50	0.19	0.20
q64: DHO visit	0.00	0.06	-0.06	-
q65: DHO call	1.15	1.38	-0.22	-0.40

Standard bias is calculated as the difference in treatment and control means, divided by the standard deviation of the treatment group.

Table 2: Balance of School Audit Variables

	Treatment Mean	Control Mean	Difference	Std. Bias
q28: separate classroom	1.00	0.92	0.08	-
q42: students with uniforms	0.35	0.34	0.00	0.02
q43: students with books	0.53	0.48	0.05	0.19
q44: students with writing utensil	0.46	0.45	0.01	0.04
q31: teacher present	0.71	0.54	0.17	0.40
q41: student absenteeism	0.56	0.47	0.09	0.30
q47: teacher tenure	415.94	335.58	80.36	0.21
q48: teacher education	2.90	2.50	0.40	0.13

Standard bias is calculated as the difference in treatment and control means, divided by the standard deviation of the treatment group.

ANNEX I: BALANCE TESTS

The issue of balance between treatment and control groups was addressed in section 4. This section provided exploratory graphics and descriptive statistics to address this issue. This appendix addresses the question in a more rigorous way, by testing the hypothesis that treatment and control groups are statistically different from one another.

Tables 3 and 4 show Wilcoxon Signed Rank tests for P3 and P4 classroom variables. This test evaluates the hypothesis that μ , or the difference between ranked pairs, is significantly different than 0. Upper and lower confidence bounds for μ are estimated at the 95% confidence level and included in the table. P-values above the 0.05 threshold confirm the hypothesis that there is no significant difference between the two distributions. The Wilcoxon Signed-Rank test was chosen here because it does not rely on normality assumptions inherent to more common test like the student's t. As is shown, for Tables 3 and 4, all of the variables have p-values significantly greater than 0. Tables 5 and 6 show balance test for school level variables. Here, only the number of toilets demonstrates significant differences across treatment and control groups. Table 7 shows similar test for HC variables. Here, none of the variables show significant imbalances across treatment and control groups.

Table 3: Wilcoxon Signed Rank Test for P3 Outcomes

Variable	p-value	μ lower	μ upper
Uniforms	0.8582	-15.5	13.5
Writing Utensil	0.9653	-23	23.5
Books	0.843	-21.5	21
Students Present	0.7766	-0.1	0.2

Table 4: Wilcoxon Signed Rank Test for P4 Outcomes

Variable	p-value	μ lower	μ upper
Uniforms	0.472	-45	22
Writing Utensil	0.6112	-33.5	33.5
Books	0.4233	-45.5	27

Table 5: Wilcoxon Signed Rank Test for School Outcomes

Variable	p-value	μ lower	μ upper
Covered Classrooms	0.1589	-0.5	3
Classrooms with Blackboards	0.3757	-1	2.5
Number of toilets	0.03516	0.5	1.0
Permanent Housing	0.09441	-0.5	4.5
Temporary Housing	0.9375	-2.5	2.5

Table 6: Wilcoxon Signed Rank Test for School Covariates

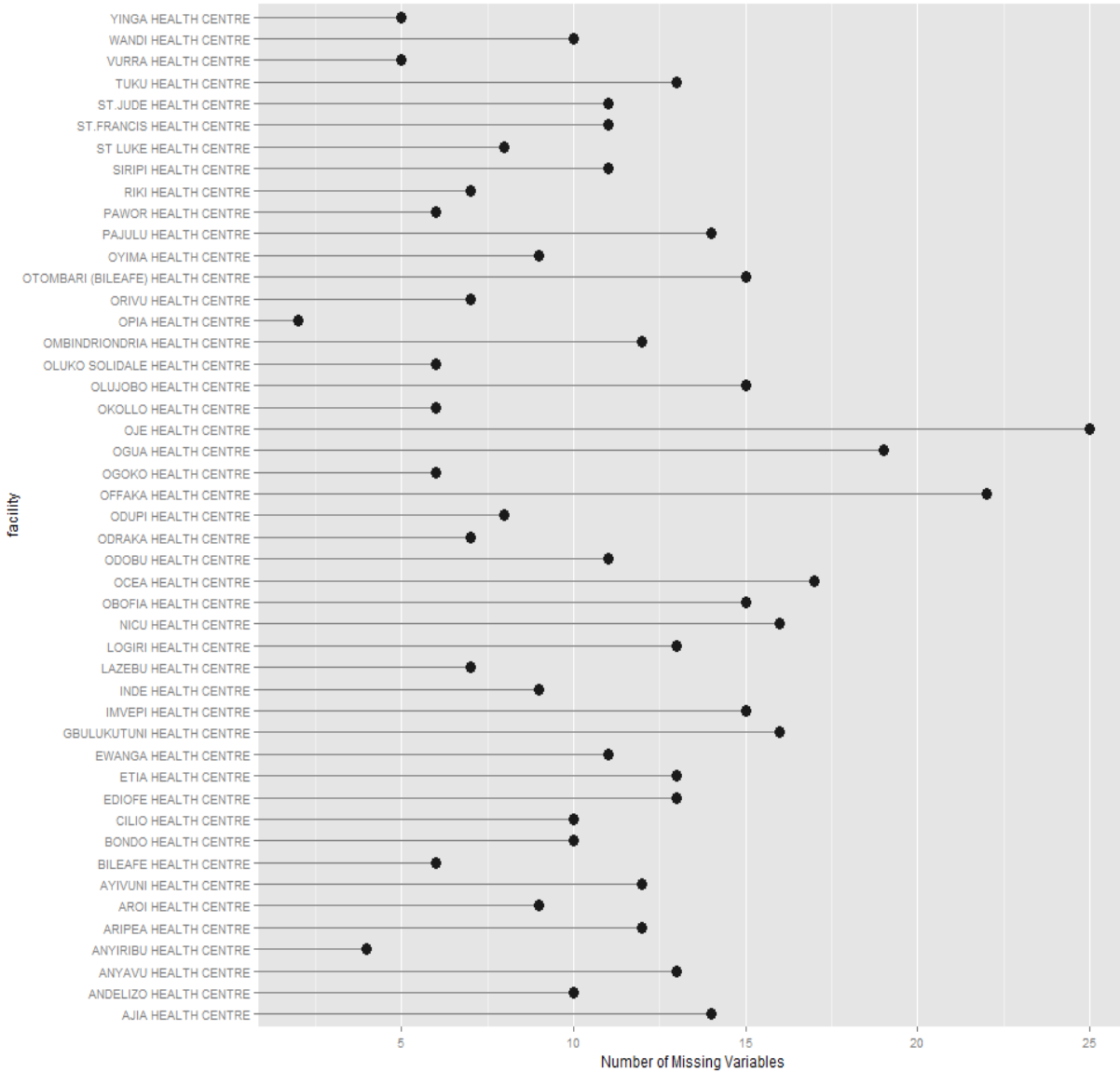
Variable	p-value	μ lower	μ upper
Teacher Education (P3)	0.5	-2	3
Teacher Education (P4)	0.75	-1	2
Number of Streams	0.2788	-1.5	3
Number of Classrooms	0.1107	-0.5	3
Total Pupils	0.2439	-60	238
DEO Visit	0.625	-1	1
DEO Calls	0.6279	-0.5	1

Table 7: Wilcoxon Signed Rank Test for P3 Outcomes

Variable	p-value	μ lower	μ upper
Absenteeism	0.6406	-0.3	0.5
Antimalarials	0.6211	-18.5	14.0
Deproprovera	0.5	-30	90
Suphadoxine	0.375	-50	6
Measles Vaccine	0.11	-1.0	21.5
ORS	0.62	-21	21
Waiting time	0.08	-1.3	0.1
Number of toilets	0.66	-2.5	1.5
Outreach Meetings	0.05	-1.5	0
Staff Meetings	0.53	-1.5	1
HUMC	0.13	-2	0.5

ANNEX 2: NUMBER OF MISSING VARIABLES

Figure 12: Missing Variables by HC



U.S. Agency for International Development
1300 Pennsylvania Avenue, NW
Washington, DC 20523